Bivariate distributions with given marginals with an application to wave climate description

G. A. Athanassoulis, E. K. Skarsoulis & K. A. Belibassakis
Department of Naval Architecture and Marine Engineering, National Technical University of Athens, PO Box 64070, 157 10 Zografos, Greece

(Received 2 September 1993; revised version received and accepted 25 October 1993)

The class of bivariate probability distributions with given (prespecified) marginals is studied, and a special member of this class, the Plackett model, is applied to represent the joint probability distribution of significant wave height \(H_s\) and mean zero-upcrossing period \(T_{02}\). The distinctive features of the bivariate Plackett model are: (i) it accepts any kind of univariate probability models as marginals, (ii) it can accurately model the degree of correlation between \(H_s\) and \(T_{02}\), and (iii) it permits an easy and reliable estimation of parameters. Applications of the Plackett model to the description of \((H_s, T_{02})\) statistics for certain sea locations illustrate its overall performance and flexibility. The possibility of extending this approach to the multivariate case is discussed.

1 INTRODUCTION

For the probability-based design and assessment of marine structures interacting with sea waves, a reliable knowledge of the long-term wave climate is required. Wave climatic data are commonly presented in the form of histograms of spectral wave parameters (sea-state parameters). The severity of a sea state is usually expressed in terms of significant wave height \(H_s\) and the mean zero-upcrossing wave period \(T_{02}\). Sometimes the significant wave height alone is used for representing the sea severity, but it is highly desirable to characterize the sea condition, using at least both parameters \(H_s\) and \(T_{02}\).

From the early stages of the development of a statistical approach to wave climate, the advantage of an analytical representation of empirical distributions of data through parametric models was recognized. The compactness of analytical description, the standardization of the representation, and the filling of information gaps (due to limited amount of data), led researchers to use specific parametric models, suitable for the description of wave height and wave period statistics. Since, however, there is no theoretical evidence as regards the probability structure of sea-state parameters (at a given site and for a given season), the convenience of the analytical model and its goodness of fit to real data have been, and still remain, the guiding criteria for selecting analytical probability models, either univariate or multivariate.

As regards univariate models, Jasper proposed in 1956 the use of univariate log-normal distribution for the description of significant wave height \(H_s\). This distribution gives, in general, a satisfactory overall fit, but exhibits deviations from data for large \(H_s\) values. Nordenstrom suggested the three-parameter Weibull distribution, which, in general, describes the upper tail of \(H_s\) data better, at the cost, however, of larger deviations for small \(H_s\) values. Thus, the Weibull distribution seems more appropriate for large-value analysis (e.g. extreme value analysis), while the log-normal distribution seems more suitable for moderate-value analysis (e.g. fatigue–life analysis, estimation of the wave-energy resource, etc.). Further efforts to develop distributions suitable for the description of wave statistics were undertaken by Ochi and Andrew and Price, who suggested the use of the generalized Gamma distribution (which also contains the two-parameter Weibull distribution as special case), while Fang and Hogben introduced a modified log-normal distribution including correction for skewness, and giving — in comparison with the standard log-normal distribution — a better description of the upper tail, especially for cumulative-probability values above 0.95. All these models have found broad use in the estimation of long-term climatic parameters, such as design values, mean long-term values, variances, etc. As regards the mean zero-upcrossing period \(T_{02}\), the log-normal model has been used, to model either the marginal \(T_{02}\)-distribution \(f_{T_{02}}\) or the conditional \(T_{02}\)-distribution \(f_{T_{02}|H_s}\), for given \(H_s\).
The needs for consideration of the joint long-term statistical character of wave height and period, as, e.g. in the case of the prediction of design wave height and associated period, led to the development and use of bivariate parametric models. Ochi proposed the bivariate log-normal distribution for the description of \((H_s, T_{02})\) statistics. This distribution is a member of the translation system (obtained by applying simple transformation laws to the multinormal distribution\(^{16,17}\)), and it will be used in the present work as a reference standard, because of its significance in wave-climate description. Fang and Hogben followed a conditional-distribution approach in order to develop a bivariate description. Fang and Hogben followed a conditional-distribution approach in order to develop a bivariate description. Ochi proposed the bivariate log-normal model (including skewness factor) for the marginal \(H_s\) distribution, and the log-normal model for the conditional \(T_{02}\) distribution. This conditional-distribution approach was further elaborated by Dacuncha and Mathiesen and Bitner-Gregersen also followed the conditional-distribution approach, by considering the three-parameter Weibull model for \(f_{H_s}\) and the log-normal model for \(f_{T_{02}}\), finally obtaining a nine-parameter bivariate model. It is the above two basic approaches (log-normal model and conditional-distribution approach) that have been used, almost exclusively up to the present, for the description of joint wave height and wave period statistics.\(^{10-21}\)

A basic requirement for any bivariate model is to be able to describe both marginal data well. In this sense, the bivariate log-normal distribution is superior to the conditional-distribution models, since it allows simultaneous control on both marginals. Furthermore, the conditional-distribution approach fails to produce reliable results in the regions of the histogram where the observations are sparse, unless some kind of parametric dependence is explicitly introduced.\(^2\) On the other hand, the bivariate log-normal model is less satisfactory than the conditional-distribution models, as regards the description of the upper tail of \(H_s\) marginal data, as mentioned above.\(^3,4,5\)

In order to overcome these problems, it is desirable to develop bivariate models that are constrained to have specified marginal behaviour, retaining enough flexibility to describe adequately the dependence structure of the bivariate quantity under study. The problem of developing bivariate distributions with given marginals was first posed by Fréchet in 1951, and, accordingly, the class of solutions of this problem is called the Fréchet class. It is the aim of this work to present this point of view for constructing bivariate probability models, and to illustrate its usefulness, by applying one of them to wave-climate description. The possibility of extending this approach to the multivariate case is also discussed.

In Section 2 a review of the basic theory of the Fréchet class is presented, with emphasis on aspects and techniques leading to explicit constructions of bivariate (multivariate) families with given marginals. A special member of the Fréchet class, the Plackett model\(^{17,23}\), is thoroughly studied (Section 3), and it is chosen to represent the bivariate distribution of \(H_s\) and \(T_{02}\). The problem of parameter estimation is considered in Section 4, and the question of goodness-of-fit is discussed in Section 5. Applications of the proposed model to describe the joint probability distribution of \((H_s, T_{02})\) for two sea locations (Station India and Station Tromsøflaket) is presented in Section 6. Finally, the main features of the proposed model are discussed in Section 7.

### 2 BIVARIATE PROBABILITY MODELS WITH GIVEN MARGINALS: A SURVEY

To fix ideas and to introduce the necessary notation and the underlying assumptions, we start with a definition.

**Definition:** A bivariate probability-distribution model on the rectangle \(I_1 \times I_2 = [a_1, b_1] \times [a_2, b_2]\) is any explicit function \(F(x_1, x_2; \theta) = F(x_1, x_2; \theta_1, \theta_2)\) of the vector \(X = (x_1, x_2) \in I_1 \times I_2\), also containing a number of parameters \(\theta = (\theta_1, \theta_2, \ldots, \theta_K)\), that satisfies the conditions:

- **(A)** \(F(b_1, b_2; \theta) = 1, \text{ for all } \theta \in \Theta\)
- **(B)** \(F(a_1, x_2; \theta) = F(x_1, a_2; \theta) = 0, \text{ for all } x_1 \in I_1, i = 1, 2, \text{ and } \theta \in \Theta,\)
- **(C)** For any two pairs \(x_1 < x_1'\) from \(I_1\), and \(x_2 < x_2'\) from \(I_2\), the following inequality holds true:

\[
F(x_1; x_2; \theta) - F(x_1, x_2; \theta) - F(x_1', x_2; \theta) + F(x_1, x_2'; \theta) \geq 0, \text{ for all } \theta \in \Theta
\]

Here, \(\Theta\) is the set of all admissible values of the parameter \(\theta\), called also parameter space. Except for the three conditions listed above, there is also a fourth condition imposed on \(F(x_1, x_2; \theta)\) for mathematical definiteness, namely that \(F(x_1, x_2; \theta)\) is continuous on the right (or on the left) with respect to each of its arguments \(x_1\) and \(x_2\) (see, e.g. Ref. 24, Section 28, or Ref. 25, Section 2.2). Since, however, this is only a matter of convention, and our main interest lies in continuous distributions, we do not pay any special attention to it here.

Under conditions (A)-(C) the function \(F(x_1, x_2; \theta)\) can serve as the distribution function of a probability measure on \(I_1 \times I_2\), that is, as the distribution function of some bivariate random vector \(X = (X_1, X_2)\), which takes values on \(I_1 \times I_2\). (For a proof, see, e.g. Tucker, \(^25\) Section 2.2). When this fact is to be emphasised, we shall also use the notation \(F_{X}(x_1, x_2; \theta)\).

Given a bivariate probability-distribution model

\(\star\) The intervals \(I_i, i = 1, 2,\) may be finite or infinite. Accordingly, \(a_i\) are finite or \(-\infty\), and \(b_i\) are finite or \(+\infty\). Three standard cases are \(I_i = R, I_i = R^+\), \(I_i = [0, 1]\).
Bivariate distributions with given marginals and wave climate

\[ F(x_1, x_2; \theta) \]

its univariate marginals are defined by

\[ F_1(x_1; \theta_1) = F(x_1, b_2; \theta), \quad \theta_1 \in \Theta_1, \]

and \[ F_2(x_2; \theta_2) = F(b_1, x_2; \theta), \quad \theta_2 \in \Theta_2 \]

and they can also be interpreted as the distribution functions of the marginal random variables \( X_1 \) and \( X_2 \), respectively. The parameter vector \( \theta_i, i = 1, 2 \), appearing in the \( x_i \)-univariate marginal, contains only a subset of the parameters appearing in the bivariate model \( F(x_1, x_2; \theta) \). In general, \( \theta \) contains more parameters than \( \theta_1 \) and \( \theta_2 \), taken together. The marginal parameter spaces \( \Theta_1 \) and \( \Theta_2 \) may or may not intersect. In the latter case, i.e. when \( \Theta_1 \cap \Theta_2 = \phi \), the parameter space \( \Theta \) can be decomposed into three mutually disjoint sets \( \Theta_1, \Theta_2 \) and \( \Theta_{12} \), which are associated with the \( x_1 \)-marginal, \( x_2 \)-marginal and their correlation structure, respectively. In practical applications, such a situation greatly facilitates the estimation procedure (see Section 4).

If the bivariate function \( F(x_1, x_2; \theta) \) is differentiable (or, more exactly, absolutely continuous, Tucker, 22 p. 28), then it can be derived by a uniquely defined density function \( f(x_1, x_2; \theta) \) by means of the equation:

\[ F(x_1, x_2; \theta) = \int_{l_1}^{l_2} \int_{l_2}^{l_2} f(x_1', x_2'; \theta) \, dx_1 \, dx_2 \]

Such a function \( f(x_1, x_2; \theta) \) will be called a bivariate probability-density model. The characteristic properties of any probability-density model are

(a) \[ \int_{l_1}^{l_2} \int_{l_1}^{l_2} f(x_1, x_2; \theta) \, dx_1 \, dx_2 = 1, \text{ for all } \theta \in \Theta, \]

(b) \[ f(x_1, x_2; \theta) \geq 0, \text{ for all } \theta \in \Theta, \]

All the above discussion concerning the parameter space \( \Theta \) and the univariate marginals also pertains to the bivariate probability-density model. The \( x_i \)-univariate density marginals will be denoted by \( f_i(x_i; \theta_i) \), \( i = 1, 2 \). The upper- and lower-case letter convention for denoting pairs of distribution and density functions, respectively, will be constantly used throughout the paper. As a rule, in the sequel, the term probability model will be used without any further specification (distribution or density) and, in general, it will refer to both cases collectively.

Let us denote by \( \mathcal{F}(F_1(x_1), F_2(x_2)) \) or, more simply, by \( \mathcal{F}(F_1, F_2) \), the class of all bivariate probability models each of which has marginal distributions \( F_1(x_1) \) and \( F_2(x_2) \). \( \mathcal{F}(F_1, F_2) \) will be called the Fréchet class generated by \( F_1(x_1), F_2(x_2) \). In general, \( \mathcal{F}(F_1, F_2) \) contains differentiable and nondifferentiable probability-distribution models. The subclass of \( \mathcal{F}(F_1, F_2) \), which contains all probability models with differentiable distribution functions, will be denoted by \( \mathcal{F}_d(F_1, F_2) \). The

The problem that will be discussed in this section can be stated, in a somewhat vague way, as follows:

Given the univariate distributions \( F_i(x_i), i = 1, 2 \), study the structure of the Fréchet class \( \mathcal{F}(F_1, F_2) \) and give explicit representations of (all or some families of) the bivariate probability models \( F(x_1, x_2) \) from \( \mathcal{F}(F_1, F_2) \) and \( \mathcal{F}_d(F_1, F_2) \).

This problem has been first posed and studied, to some extent, by Fréchet.22,26,27 Since then, it has been extensively studied, mainly in the mathematical literature (see e.g. the books by Mardia, 28 Johnson and Kotz, 16 Johnson, 17 and the references below), without disseminating, until now, to applied sciences and engineering.

As can be inferred from the statement of the problem given above, there are three main directions of investigation:

(i) to study the general properties (the structure) of the Fréchet class, \( \mathcal{F}(F_1, F_2) \);

(ii) to obtain general representation theorems encompassing all the members of the class \( \mathcal{F}(F_1, F_2) \) or \( \mathcal{F}_d(F_1, F_2) \);

(iii) to obtain simple closed-form expressions defining easy-to-use and well-behaved families of the class \( \mathcal{F}(F_1, F_2) \) or \( \mathcal{F}_d(F_1, F_2) \).

Some fundamental aspects in these directions will be surveyed below.

2.1 Some general properties of the Fréchet class

Let us denote by \( U(y) = y \) the uniform distribution on \([0, 1]\), and let the random variable \( Y \) be distributed in accordance with \( U(y) \); then, given any distribution function \( F(x) \), the random variable \( X \), defined by means of the transformation \( X = F^{-1}(Y) \), is distributed in accordance with \( F(x) \). Conversely, if the variable \( X \) is distributed in accordance with \( F(x) \), then the variable \( Y = F(X) \) is uniformly distributed on \([0, 1]\) (see e.g. Papoulis, 29 p. 101). This well-known property, which is sometimes called the probability transformation, can be applied twice to establish the following property:

P1. For any given pair \( F_i(x_i), i = 1, 2 \) of univariate distributions, the study of the Fréchet class \( \mathcal{F}(F_1, F_2) \) with uniform marginals \( F(U_1, U_2) \).

This easily proven property has very important consequences. It shows that the Fréchet class is essentially unique, and its study actually coincides with the study of the correlation structure of bivariate distributions with uniform marginals. Property P1 also shows that any (general or special) representation of families \( F(x_1, x_2; \theta) \) from the Fréchet class can be used in combination with any type of marginal distribution, and that the parameter space \( \Theta \) of \( F(x_1, x_2; \theta) \) can be decomposed to \( \Theta_1, \Theta_2 \) and \( \Theta_{12} \), as discussed
previously. From the point of view of model construction, Property P1 can be restated as follows:

**P2.** Let \( F_i(x_i), G_i(z_i), i = 1, 2, \) be two pairs of distribution functions. If the probability model \( F(x_1, x_2) \) belongs to the class \( \mathcal{F}(F_1, F_2) \), then the probability model \( G(z_1, z_2) = F(F_1(G_1(z_1)), F_2(G_2(z_2))) \) belongs to the class \( \mathcal{G}(G_1, G_2) \). In particular, if \( F(x_1, x_2) \) belongs to \( \mathcal{F}(F_1, F_2) \), then \( G(z_1, z_2) = F(G_1(z_1), G_2(z_2)) \) belongs to \( \mathcal{G}(G_1, G_2) \), and the corresponding density function, if it exists, is given by

\[
g(z_1, z_2) = f(G_1(z_1), G_2(z_2))g_1(z_1)g_2(z_2) \quad (3)
\]

This property permits us to obtain closed-form expressions for families of probability models, starting from known bivariate distributions. If the initial model is the bivariate normal, the family of models obtained is known as a translation family or translation system. It was first introduced by Edgeworth and has been extensively studied by van Uven, Johnson and others (see also Ref. 17, Chap. 5, where three translation systems are thoroughly studied, including the classical log-normal bivariate model). The possibility of starting from more general initial models \( F(x_1, x_2) \) was first discussed by Natarajan, who also noticed that this construction is straightforwardly generalized to the multivariate case.

Another simple, but useful property of the Fréchet class, from the point of view of model construction, is the following:

**P3.** If \( F_i(x_i), i = 1, 2, \ldots \), belong to the Fréchet class \( \mathcal{F}(F_1, F_2) \), and \( \epsilon_i \geq 0, \sum\epsilon_i = 1 \), then the combined probability model \( F(x_1, x_2) = \sum\epsilon_i F_i(x_1, x_2) \) (the probabilistic mixture) also belongs to \( \mathcal{F}(F_1, F_2) \).

Furthermore, as was pointed out by Fréchet, independently by Hoeffding (see also Ref. 38), the knowledge of the marginals \( F_i(x_i), F_2(x_2) \) implies some bounds on the joint distribution \( F(x_1, x_2) \):

**P4.** Any member of the Fréchet class \( \mathcal{F}(F_1, F_2) \) satisfies the inequalities (Fréchet bounds)

\[
F^{(-1)}(x_1, x_2) \leq F(x_1, x_2) \leq F^{(+1)}(x_1, x_2) \quad (4)
\]

where

\[
F^{(-1)}(x_1, x_2) = \frac{1}{2}[F_1(x_1) + F_2(x_2) - 1 + \sqrt{(F_1(x_1) + F_2(x_2) - 1)^2 + 1}] \quad (5)
\]

and

\[
F^{(+1)}(x_1, x_2) = \frac{1}{2}[F_1(x_1) + F_2(x_2) - \sqrt{(F_1(x_1) + F_2(x_2) - 1)^2 + 1}] \quad (6)
\]

Both \( F^{(-1)}(x_1, x_2) \) and \( F^{(+1)}(x_1, x_2) \) are distribution functions from the same Fréchet class \( \mathcal{F}(F_1, F_2) \). However, they are not differentiable, so they cannot be derived from density functions.

Note that the correlation coefficients of \( F^{(-1)}(x_1, x_2) \) and \( F^{(+1)}(x_1, x_2) \) are \( \rho^{(-1)} = -1 \) and \( \rho^{(+1)} = +1 \), respectively. That is, eqns (5) and (6) give explicit expressions of the members of the Fréchet class \( \mathcal{F}(F_1, F_2) \) which are fully correlated (negatively and positively, respectively).

Fréchet bounds have played an important role in developing parametric families of probability models from \( \mathcal{F}(F_1, F_2) \), since they have served as the proper limiting forms of the models, for the limiting values of the parameters involved (see also Section 3.1).

### 2.2 General representation theorems of the Fréchet class

General representations of bivariate distributions \( F(x_1, x_2) \in \mathcal{F}(F_1, F_2) \), in terms of the marginals \( F_i(x_i), i = 1, 2, \) and a set of parameters describing the correlation structure of \( X = (X_1, X_2) \), have been given in the form of an infinite series. At least three different directions have been developed, generalizing the corresponding results from the univariate case: (i) a bivariate Fourier-series expansion, developed by Abazaliev; generalizing the univariate Fourier-series approach introduced by Sarmanov and independently by Posten and Woods; (ii) a bivariate Gram–Charlier expansion, developed by Sarmanov, generalizing the univariate Gram–Charlier series (see, e.g. Ref. 45, Sections 8.6 and 8.7 or Ref. 46, Chapt. 16); and (iii) a more general bivariate series representation, introduced by Lancaster, expressing \( F(x_1, x_2) \) in terms of two sequences of functions orthogonal with respect to the marginal distributions. The latter representation (Lancaster series expansion) has the form:

\[
f(x_1, x_2) = f_1(x_1)f_2(x_2) \left[ 1 + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \rho_{ij} q_{1,i}(x_1)q_{2,j}(x_2) \right] \quad (7)
\]

where the systems \( \{q_{1,i}(x_1), i = 1, 2, \ldots \} \) and \( \{q_{2,j}(x_2), j = 1, 2, \ldots \} \) are complete orthonormal systems with weights \( f_1(x_1) \) and \( f_2(x_2) \), respectively. Moreover, for eqn (7) to be valid, it is necessary that

\[
\int \int \left( \frac{f_1(x_1, x_2)}{f_1(x_1)f_2(x_2)} \right) \text{d}x_1 \text{d}x_2 < \infty \quad (8)
\]

a restriction that excludes fully correlated distributions (\( \rho = \pm 1 \), but is not very restrictive otherwise. Thus, series expansion (eqn (7)) is a general representation for the subclass of \( \mathcal{F}(F_1, F_2) \), whose members satisfy eqn (8). Extension of eqn (7) to more general bivariate distributions, not satisfying eqn (8), have been discussed by Lancaster and other authors. Note, however, that if eqn (8) is not satisfied, a series expansion of the form of eqn (7) may not exist.
All the above representations can be generalized to the multivariate case. This, in conjunction with their great flexibility in representing the correlation structure, suggests that such representations (especially (i) and (iii)) may offer very promising probability models for bivariate and multivariate environmental quantities. The main difficulties to be resolved concern the constraint of positivity of the truncated series (see conditions (C) and (B), above) and the use of efficient and well-behaved estimators for the coefficients related to the dependence structure (see, e.g. Ref. 50). These aspects will be studied in a forthcoming paper.

2.3 Special models

Because of the complexity of the general representations (which, on the other hand, were not known before the 1960s), a lot of special models of bivariate distributions with given marginals have been developed and studied in statistical literature. 23,30,31–35 Extended surveys of this subject have been presented by Mardia, 28 Conway, 56 Johnson 17 and Barnett. 57 Two of them are briefly discussed in this section, and one of them, the Plackett distribution, is thoroughly studied and applied to wave-climate description in Section 3.

An interesting and extensively studied bivariate model from the class \( \mathcal{F}(F_1, F_2) \), first introduced by Morgenstern 16,17,51 is given by the equation:

\[
\frac{f_{X_1X_2}(x_1, x_2; \mu)}{f_1(x_1)f_2(x_2)} = \left[ 1 + \mu(2F_1(x_1) - 1) \right] \left[ 1 + (2F_2(x_2) - 1) \right]
\]

The dependence structure between \( X_1 \) and \( X_2 \) is controlled by the parameter \( \mu \). However, for \( f_{X_1X_2}(x_1, x_2; \mu) \) to be positive, \( \mu \) has to lie in the interval \([-1, 1]\). This constraint restricts the possible values of correlation coefficient \( \rho \) to \(-\frac{1}{2} \leq \rho \leq \frac{1}{2} \), a fact that limits the applicability of Morgenstern distribution only to cases where the dependence is rather weak. The Morgenstern model has also been extended to the multivariate case. 17,58–60 However, again the corresponding multivariate distribution can be applied only to weakly correlated data.

Attempts to alleviate the weak-dependence restriction have led either to a generalized Morgenstern model 54 or to entirely different models. In this paper we shall focus our attention on such a model, introduced by Plackett, 23 whose density function in the case of uniform marginals has the form:

\[
f_{X_1X_2}(x_1, x_2; \psi) = \frac{\psi[(\psi - 1)(x_1 + x_2 - 2x_1x_2) + 1]}{\{1 + (x_1 + x_2)(\psi - 1)\}^2 - 4\psi(\psi - 1)x_1x_2}^{3/2}
\]

where \( (x_1, x_2) \in [0, 1]^2 \) and the parameter \( \psi \in \mathbb{R}^+ \). The non-negative parameter \( \psi \) governs the dependence between the random variables \( Y_1 \) and \( Y_2 \), and it is called a correlation parameter. The main advantages of the Plackett model are: it covers the full range from the lower to the upper Fréchet bound (Section 2.1), including the case of independence. As a consequence, the Plackett model can describe data with arbitrary degrees of correlation, in the full possible range of correlation coefficients \(-1 < \rho < 1\). Furthermore, the correlation parameter \( \psi \) can be interpreted as a measure of association in a fourfold contingency table (see Section 3.1), a property which enables a simple first estimation of \( \psi \), in closed form, directly from the data, as described in Section 4. A drawback of the Plackett distribution is its restriction to the bivariate case. Attempts have been undertaken for its extension to three dimensions; 56 however, the results are not yet complete. A detailed description of the theory of the Plackett model, in its general form \( f_{X_1X_2}(x_1, x_2; \theta) \in \mathcal{F}(F_1, F_2) \), is given below.

3 THEORY OF THE PLACKETT MODEL AND APPLICATION TO WAVE-CLIMATE DESCRIPTION

3.1 Theory of the Plackett model

The general form of the Plackett density with given marginal distributions \( F_i(x_i) = F_i(x_i; \theta_i) \) and densities \( f_i(x_i) = f_i(x_i; \theta_i) \), \( i = 1, 2 \), can be obtained by applying the transformation \( y_i = F_i(x_i; \theta_i) \), \( i = 1, 2 \), to the uniform Plackett density (eqn (10)) (see Property P2, eqn (3), in Section 2.1). Thus

\[
f_{X_1X_2}(x_1, x_2; \theta) = \left( \begin{array}{c}
\psi[(\psi - 1)(F_1(x_1) + F_2(x_2) - 2F_1(x_1)F_2(x_2)) + 1]F_1(x_1)F_2(x_2)
\end{array}\right)^{3/2}
\]

The parameter vector \( \theta = (\theta_1, \theta_2, \psi) \) takes values in the factorized parameter space \( \Theta_1 \times \Theta_2 \times \Theta_{12} \), \( \Theta_{12} = (0, \infty) \). The corresponding distribution function can be found by integrating eqn (11):

\[
F_{X_1X_2}(x_1, x_2; \theta) = \begin{cases}
\frac{S(x_1, x_2) - 4\psi(\psi - 1)F_1(x_1)F_2(x_2)}{2(\psi - 1)}, & \text{for } \psi \neq 1 \\
F_1(x_1)F_2(x_2), & \text{for } \psi = 1
\end{cases}
\]
where \( S(x_1, x_2) = 1 + (\psi - 1)|F_1(x_1) + F_2(x_2)|. \) From this expression, it can be easily seen that
\[
\lim_{\psi \to -\infty} F_\psi^P(x_1, x_2; \tilde{\theta}) = F_{\tilde{\theta}}^{(-1)}(x_1, x_2)
\]
and
\[
\lim_{\psi \to \infty} F_\psi^P(x_1, x_2; \tilde{\theta}) = F_{\tilde{\theta}}^{(+1)}(x_1, x_2)
\] (13)
where \( F_{\tilde{\theta}}^{(-1)}(x_1, x_2) \) and \( F_{\tilde{\theta}}^{(+1)}(x_1, x_2) \) are the Fréchet limiting distributions defined by eqns (5) and (6). That is, \( F_\psi^P(x_1, x_2; \tilde{\theta}) \) tends to the lower Fréchet bound as \( \psi \to 0 \) and to the upper Fréchet bound as \( \psi \to \infty \), while \( \psi = 1 \) corresponds to the case of independence. Furthermore, \( \psi \) values in the interval \((0, 1)\) induce a negative correlation, while \( \psi \) values in the interval \((1, \infty)\) induce a positive correlation. A complete series of figures, showing the correlation structure of the Plackett density for various values of \( \psi \), are presented by Johnson.\(^{17}\)

From eqn (12) we obtain:\(^{17}\)
\[
\psi = \frac{F^P_{\tilde{\theta}}(x_1, x_2; \tilde{\theta})}{[F_1(x_1) - F^P_{\tilde{\theta}}(x_1, x_2; \tilde{\theta})][F_2(x_2) - F^P_{\tilde{\theta}}(x_1, x_2; \tilde{\theta})]} - P_2 P_3
\]
which is actually the generating relation for the Plackett distribution.\(^{2}\) By expressing the distribution function on the right-hand side of eqn (14) in terms of probabilities, we obtain:
\[
\psi = \frac{Pr[X_1 \leq x_1, X_2 \leq x_2]Pr[X_1 > x_1, X_2 > x_2]}{Pr[X_1 \leq x_1, X_2 > x_2]Pr[X_1 > x_1, X_2 \leq x_2]} = P_1 P_2
\]
That is, \( \psi \) is the measure of association of the fourfold contingency table:

<table>
<thead>
<tr>
<th></th>
<th>( P_2 )</th>
<th>( P_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_1 )</td>
<td>( P_3 )</td>
<td></td>
</tr>
</tbody>
</table>

where \( p_i, i = 1, 2, 3, 4 \), are the probabilities of the four quadrants, defined by two lines parallel to the axes and intersecting each other at the point \((x_1, x_2)\), i.e.

- \( p_1 = Pr[X_1 \leq x_1, X_2 \leq x_2] \)
- \( p_2 = Pr[X_1 \leq x_1, X_2 > x_2] \)
- \( p_3 = Pr[X_1 > x_1, X_2 \leq x_2] \)
- \( p_4 = Pr[X_1 > x_1, X_2 > x_2] \).

For an arbitrary bivariate distribution having \( F_1 \) and \( F_2 \) as marginals, the above measure of association is, in general, a function of \((x_1, x_2)\). Plackett’s family (eqns (11) and (12)) is the class of bivariate distributions from \( F(F_1, F_2) \) having the measure of association (eqn (15)) constant for all values \((x_1, x_2)\). A straightforward consequence of this property is a simple closed-form formula for a first estimation of \( \psi \), which does not involve the parameters \( \hat{\theta}_1, \hat{\theta}_2 \) of the marginal distributions (see Section 4.2).

### 3.2 Marginal models

As mentioned in Section 1, on the basis of previous experience,\(^{1,4,6,12,21}\) Weibull and log-normal distributions are considered prime candidates for representing the univariate probability density functions of the significant wave height \( H_s \) and mean zero-upcrossing period \( T_{02} \). The gamma distribution\(^{7,8}\) is also used for comparison purposes. The probability density functions corresponding to the log-normal, Weibull and Gamma distributions are

\[
f_L(x; \bar{\theta}_L) = \frac{1}{\sqrt{2\pi x^2}} \exp \left\{ -\frac{[\ln(x) - \alpha]^2}{2\beta^2} \right\}, \quad x > 0,
\]

\[
\bar{\theta}_L = (\alpha, \beta) \in \mathbb{R} \times \mathbb{R}^+
\]

\[
f_w(x; \bar{\theta}_w) = \alpha \beta (x - \gamma)^{\beta - 1} \exp\{-(\alpha(x - \gamma))\}, \quad x > \gamma,
\]

\[
\bar{\theta}_w = (\alpha, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+
\]

\[
f_r(x; \bar{\theta}_r) = \frac{x^{\alpha - 1}}{\beta^\Gamma(\alpha)} \exp\{\frac{-x}{\beta}\}, \quad x > 0,
\]

\[
\bar{\theta}_r = (\alpha, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+
\]

The threshold parameter \( \gamma \) of the Weibull distribution (eqn (17)) is regarded as a constant known in advance. Thus, all three distributions are two-parameter distributions. The corresponding cumulative distribution functions will be denoted by \( F_L(x; \bar{\theta}_L), F_w(x; \bar{\theta}_w), F_r(x; \bar{\theta}_r) \), respectively.

In applying the Plackett model to represent \((H_s, T_{02})\) distribution, the marginals \( F_1 \) and \( F_2 \) will be taken from the set \( F = \{F_L, F_w, F_r\} \). However, any other univariate probability model (e.g. the skewed log-normal distribution) could be equally well considered in combination with the bivariate Plackett distribution (see Property 5, Section 2.1). Thus, any improvement in modelling the univariate marginal distributions of \( H_s \) and \( T_{02} \) can be directly exploited for the description of the joint \((H_s, T_{02})\) distribution.

Since the set \( F \) contains two-parameter distributions, the resulting Plackett model is a five-parameter model \( F_\psi^P(x_1, x_2; \tilde{\theta}) \), with parameter vector \( \tilde{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \psi) \), where \( \hat{\theta}_i = (\alpha_i, \beta_i) \in \{\hat{\theta}_1, \hat{\theta}_w, \hat{\theta}_r\}, i = 1, 2 \).

An example of the Plackett distribution before and after the transformation \( y_i = F_i(x_i; \hat{\theta}_i) \), \( i = 1, 2 \), for two values of the correlation parameter \( \psi \) is shown in Fig. 1. In particular, Figs 1(a) and (b) show the density \( f_\psi^P(y_1, y_2; \psi) \), with uniform marginals, for \( \psi = 5 \) and \( \psi = 20 \), respectively. It is seen that as \( \psi \) increases the distribution gradually concentrates around the line \( y_1 = y_2 \). Figures 1(c) and (d) show the corresponding transformed Plackett densities \( f_\psi^P(x_1, x_2; \tilde{\theta}) \), obtained by using log-normal marginal densities (eqn (16)). The parameters of the log-normal marginals used in this example are \( \alpha_1 = -0.5, \beta_1 = 0.2 \), \( \alpha_2 = -1.5 \) and \( \beta_2 = 0.6 \). The value of the correlation parameter \( \psi \) is taken equal to 5 in Fig. 1(c) and equal to 20 in Fig. 1(d). The increase of correlation between \( X_1 \) and \( X_2 \) as \( \psi \) increases is again evident from the figure.
Bivariate distributions with given marginals and wave climate

Fig. 1. Plackett density with uniform (a,b) and log-normal (c,d) marginals. (Parameters of log-normal marginals: $\alpha_1 = -0.5, \beta_1 = 0.2, \alpha_2 = -1.5, \beta_2 = 0.6.$) (a) $\psi = 5$; (b) $\psi = 20$; (c) $\psi = 5$; (d) $\psi = 20$.

4 ESTIMATION OF PARAMETERS

The following strategy is applied to estimating the parameters $\theta = (\theta_1, \theta_2, \theta_{12})$ of a bivariate probability model $F(x_1, x_2; \theta) \in F_1(x_1; \theta_1), F_2(x_2; \theta_2)$ from data. First, the parameters $\theta_1, \theta_2$ of the univariate marginal models $F_1(x_1; \theta_1), F_2(x_2; \theta_2)$ are estimated from the corresponding univariate marginal data, using, e.g. the maximum likelihood (ML) method $^{61,62}$ (applied to the univariate marginals). Second, the correlation parameters $\theta_{12}$ are estimated from the bivariate data, keeping $\theta_1, \theta_2$ fixed, and using again, e.g. the ML method. In this way, in estimating the marginal parameters of the bivariate model it is possible to use not only bivariate data (i.e. pairs $(x_1^{(n)}, x_2^{(n)})$), but also additional univariate data, if they are available. Note that, in environmental applications, it is quite common to have at our disposal a significantly greater amount of univariate, than bivariate, data.

Finally, it may be possible (or desirable) to proceed to a global re-estimation of the complete parameter vector $\theta = (\theta_1, \theta_2, \theta_{12})$, e.g. by applying the ML method to the bivariate model. It should be stressed, however, that the first two steps, described above, are most important, since they usually provide a very satisfactory solution to the estimation problem. On the other hand, this solution can be used as a first approximation of the sought-for parameters in the case where we are proceeding to the third step of global re-estimation of the complete parameter vector $\theta = (\theta_1, \theta_2, \theta_{12})$. Let it be noted that this third step may not be possible (or
useful) in the case where the amount of bivariate data is small.

4.1 Estimation of marginal parameters from marginal data

In the applications studied in this paper (see Section 6), the ML method is used for estimating the parameters \( \tilde{\theta}_1 \) and \( \tilde{\theta}_2 \) of the univariate marginal models (first step, above). Using the log-likelihood function:

\[
l_N(s_i; \tilde{\theta}_i) = \ln \left( \prod_{n=1}^{N} f_i(x_i^{(n)}; \tilde{\theta}_i) \right), \quad i = 1, 2
\]

where \( s_i = (x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(N_i)}) \), \( i = 1, 2 \), is the corresponding marginal sample, consisting of \( N_i \) univariate observations, we find the following ML estimators \( \tilde{\theta}_i = (\alpha_i, \beta_i) \), \( i = 1, 2 \), for the parameters \( (\alpha, \beta) \) of the log-normal, Weibull and Gamma distributions

\begin{align*}
\text{Log-normal distribution} & \\
\alpha^* &= \left( \frac{1}{N} \sum_{n=1}^{N} \ln(x^{(n)}) \right), \quad \beta^* = \left( \frac{1}{N} \sum_{n=1}^{N} [\ln(x^{(n)}) - \alpha^*]^2 \right)^{1/2} \\
\text{Weibull distribution} & \\
\alpha^* &= \left( \frac{1}{N} \sum_{n=1}^{N} [\ln(x^{(n)} - \gamma)]^{\beta^*} \right)^{-1} \\
\text{Gamma distribution} & \\
\psi(\alpha^*) - \ln(\alpha^*) &= \frac{1}{N} \sum_{n=1}^{N} \ln\left( \frac{x^{(n)}}{\bar{x}} \right), \quad \beta^* = \frac{\bar{x}}{\alpha^*}
\end{align*}

The threshold parameter \( \gamma \) is found by trials so as to ensure the best overall fit.

4.2 Estimation of the correlation parameter

For estimating the correlation parameter \( \tilde{\theta}_{12} \) of the bivariate model:

\[
f_{\tilde{\theta}}(x_1, x_2; \tilde{\theta}) = f_{\tilde{\theta}}(F_1(x_1; \tilde{\theta}_1), F_2(x_2; \tilde{\theta}_2); \tilde{\theta}_{12}) 
\times f_1(x_1; \tilde{\theta}_1)f_2(x_2; \tilde{\theta}_2)
\]

(see also eqns (3) and (11)), \( (\tilde{\theta}_{12} = (\psi) \) in the case of the Plackett model) we again use the ML method. The bivariate log-likelihood function, i.e. the log-likelihood function of the bivariate model \( f_{\tilde{\theta}}(x_1, x_2; \tilde{\theta}) \), can be written in the form:

\[
l^N_{\tilde{\theta}}(s; \tilde{\theta}) = l^N_{\tilde{\theta}_1}(s; \tilde{\theta}_1, \tilde{\theta}_2, \tilde{\theta}_{12}) + l^N_{\tilde{\theta}_2}(s; \tilde{\theta}_1, \tilde{\theta}_2, \tilde{\theta}_{12})
\]

where

\[
l^N_{\tilde{\theta}_1}(s; \tilde{\theta}_1, \tilde{\theta}_2, \tilde{\theta}_{12}) = \ln \left( \prod_{n=1}^{N} f_{\tilde{\theta}}(F_1(x_1^{(n)}; \tilde{\theta}_1), F_2(x_2^{(n)}; \tilde{\theta}_2); \tilde{\theta}_{12}) \right)
\]

\[
l^N_{\tilde{\theta}_2}(s; \tilde{\theta}_1, \tilde{\theta}_2, \tilde{\theta}_{12}) = \ln \left( \prod_{n=1}^{N} f_{\tilde{\theta}}(F_1(x_1^{(n)}; \tilde{\theta}_1), F_2(x_2^{(n)}; \tilde{\theta}_2); \tilde{\theta}_{12}) \right)
\]

Thus, in order to minimize the variance of \( \psi \), the point \((x_1, x_2)\) of the double dichotomy of the \((x_1, x_2)\)-plane should be chosen so as to make \( \mu_1, \mu_2, \mu_3 \) and \( \mu_4 \) to be comparable in magnitude. A remarkable feature of eqns (27) and (28) is that they do not require knowledge of the marginal parameters \( \tilde{\theta}_1 \) and \( \tilde{\theta}_2 \) in order to provide a value for the correlation parameter \( \psi \).
The estimate \( \hat{\psi} \) obtained using eqn (27) can be used as an initial approximation of \( \psi^* \) in the case where the maximization problem (eqn (26)) is solved by means of an iterative procedure.

### 4.3 Global estimation of parameters of the bivariate model

If \( N_1 = N_2 = N \) or, more generally, if the bivariate sample is sufficiently large, then it would be of interest to estimate the complete parameter vector \( \hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_{12}) \) from the bivariate sample \( s \). In this case the ML estimator of \( \hat{\theta} \), which can be called the global ML estimator of \( \hat{\theta} \), is obtained by means of the following maximization problem:

Find \( \hat{\theta} = \hat{\theta}^* \) such that

\[
\hat{\theta}^* = \arg \max_{\theta} L_{\theta}(s; \theta)
\]

where \( L_{\theta}(s; \theta) \) is the bivariate log-likelihood function (eqn (24)). The procedure of the numerical solution of the problem in eqn (29) is greatly accelerated if the estimate \( \hat{\theta}^* = (\hat{\theta}_1^*, \hat{\theta}_2^*, \hat{\theta}_{12}^*) \), obtained through the two previous steps, is used as an initial approximation of the sought-for solution \( \hat{\theta}^* \). On the other hand, because of the structure of the bivariate log-likelihood function (eqn (24)), the global maximum of \( L_{\theta}(s; \theta) \) is not expected to differ significantly from \( (\hat{\theta}_1^*, \hat{\theta}_2^*, \hat{\theta}_{12}^*) \), which was obtained by maximizing separately the 'partial' log-likelihood functions \( L_{\theta}(s_1; \hat{\theta}_1) \), \( L_{\theta}(s_2; \hat{\theta}_2) \) and \( L_{\theta}(s_1, s_2; \hat{\theta}_{12}) \), provided that \( N_1 \approx N_2 \approx N \). In fact, in all cases studied in this paper \( (N_1 = N_2 = N) \), the difference between \( \hat{\theta}^* \) and \( \hat{\theta}^{**} \) was found to be very small.

### 5 GOODNESS-OF-FIT MEASURES

For the assessment of a parametric model, a measure for testing the goodness-of-fit of the model to the sample data is needed. Two kinds of such measures are considered in this section, the \( \chi^2 \)-statistic, and a simple Euclidean distance, which has no statistical background.

Let \( \Omega = \{ \mathcal{C}_j \}_{j=1}^J \) be a finite partition of the domain of definition \( I_1 \times I_2 \) of the bivariate distribution considered, and \( \{ p_j \}_{j=1}^J \) be the cell probabilities resulting from the statistical model \( f_X(x; \theta) \):

\[
p_j = \int_{\mathcal{C}_j} f_X(x; \theta) \, dx, \quad j = 1, \ldots, J
\]

(30)

Let, furthermore, \( \pi_j = \nu_j / N \), \( j = 1, 2, \ldots, J \), be the corresponding relative frequencies resulting from a particular sample, where \( \nu_j \), \( j = 1, 2, \ldots, J \), denotes the number of observations falling in the cell \( \mathcal{C}_j \), and \( N \) is the observation count \( N = \sum_{j=1}^J \nu_j \). The \( \chi^2 \)-statistic\(^{44} \) has the form:

\[
\chi^2 = N \sum_{j=1}^J \frac{(p_j - \pi_j)^2}{p_j}
\]

(31)

Under the assumption that the sampled population obeys the probability law \( f_X(x; \theta) \), the \( \chi^2 \)-statistic follows \( \chi^2 \) distribution with \( (J - K - 1) \) degrees of freedom, where \( K \) is the number of parameters of the model (in the case of the Plackett model \( K = 5 \)).

The dependence of the \( \chi^2 \)-statistic from the observation count \( N \) has a dominant effect on the result of the \( \chi^2 \) goodness-of-fit test: a proportional uniform (over all

\[1 \text{ ft} = 30.48 \text{ cm}.\]

Table 1. \((H_s, T_{02})\) histogram of relative \( (\%\) frequencies of occurrence for Station India, based on a total of 2400 observations (After Nordenstrom\(^{4}\))

<table>
<thead>
<tr>
<th>( H_s ) (ft)(^{*})</th>
<th>( T_{02} ) (s) (central value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.5</td>
<td>7-5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>33</td>
<td>0</td>
</tr>
<tr>
<td>35</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>0</td>
</tr>
<tr>
<td>39</td>
<td>0</td>
</tr>
<tr>
<td>41</td>
<td>0</td>
</tr>
<tr>
<td>43</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>0</td>
</tr>
<tr>
<td>47</td>
<td>0</td>
</tr>
</tbody>
</table>

\[^{*} \text{1 ft} = 30.48 \text{ cm.}\]

Fig. 2. Contour plots of \((H_s, T_{02})\) empirical probability density for Station India, obtained from Table 1 by interpolation.
cells) change of the number of observations, although leaving the model parameters unaltered, will lead to a change in the value of the $\chi^2$-statistic in proportion to $N$, and thus to a possible change to the outcome of the statistical test. As an example, let us consider the fit of the bivariate log-normal distribution on $(H_5, T_{02})$ data from Station India (Table 1). The model is rejected at the 5% confidence level if the original observations ($N = 2400$) are considered. However, after uniform shrinking of the sample to $N = 1000$, the same model is accepted.

This behaviour of the $\chi^2$-statistic is explained by the very nature of the $\chi^2$ test, which aims at testing the hypothesis that a sample comes from a certain population, having probability distribution $F_\theta(x; \theta)$. The dependence of the $\chi^2$-statistic, and thus of the outcome of the $\chi^2$ test, on the observation count $N$ accounts for the fact that, if the above hypothesis is true, then the relative frequencies $\pi_j$ should converge to the probabilities $p_j$ of the cells $C_j$, as $N \to \infty$.

However, in our case there is not any evidence that the chosen and fitted probability model $F_\theta(x; \bar{\theta})$ (in fact, any such model) represents the probability structure of sea-state parameters exactly. In this case, a nonprobabilistic measure of goodness of approximation seems more appropriate than the $\chi^2$ test. The Euclidean distance between the theoretical probabilities $p_j$ and the relative frequencies $\pi_j$ has been chosen for this purpose:

$$D^2 = \sum_{j=1}^J (p_j - \pi_j)^2$$  \hspace{1cm} (32)

By using the $D^2$ measure we can compare different parametric models, as regards their closeness to a given data set.

6 NUMERICAL RESULTS

Two sets of measured wave data, one from Station India ($59^\circ N, 19^\circ W$) in the North Atlantic Ocean, and one from station Tromsøfjellet ($71^\circ 30' N, 19^\circ E$) in the Norwegian continental shelf, are considered for the numerical evaluation of the performance of the bivariate Plackett model.

6.1 Station India

Table 1 shows the $(H_5, T_{02})$-histogram for Station India. The data are based on a total of 2400 wave measurements.\(^6\)\(^6\) These data are presented in the form of contour plots (after interpolating the histogram data) in Fig. 2.

In accordance with the estimation scheme described in Section 4, the marginal distributions are considered first. The ML estimates $\bar{\theta}^*_i = (\alpha^*_i, \beta^*_i)$, $i = 1, 2$, of the marginal parameters for the three marginal models (log-normal, Weibull and Gamma, eqns (20)-(22)) are given in Table 2.\(^6\)

The fit of the marginal models to the marginal histograms of $T_{02}$ and $H_5$ is shown in Fig. 3. It is seen from this figure that the log-normal model offers the best global fit, among the three univariate models considered, for both $T_{02}$ and $H_5$; this is in accordance with earlier conclusions by Ochi.\(^3\) Further, the Gamma model for the $T_{02}$-distribution, and the Weibull model

\(^6\)Prior to estimation, the $H_5$ and $T_{02}$ data have been normalized by the maximum values $H_{5,\max}$ and $T_{02,\max}$ of the histogram, and the parameters given refer to the normalized variables. In the case of Station India $H_{5,\max} = 48\text{ ft}$ (14.63 m) and $T_{02,\max} = 14\text{ s}$, while in the case of Station Tromsøfjellet, below, the normalizing values have been taken $H_{5,\max} = 10\text{ m}$ and $T_{02,\max} = 13\text{ s}$.
Table 2. ML estimates of marginal parameters for Station India

<table>
<thead>
<tr>
<th>Univariate model</th>
<th>$T_{02}$-marginal</th>
<th>$H_s$-marginal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>$\beta_1$</td>
</tr>
<tr>
<td>Log-normal</td>
<td>-0.4019</td>
<td>0.1384</td>
</tr>
<tr>
<td>Gamma</td>
<td>52.3690</td>
<td>0.0187</td>
</tr>
<tr>
<td>Weibull</td>
<td>36.9930</td>
<td>2.8155</td>
</tr>
</tbody>
</table>

$^a\gamma_{T_{02}} = 5.992\, s; \gamma_{H_s} = 2.616\, ft (c. 79.74\, cm)$

Table 3. ML estimates $\psi^*$ of the correlation parameter of the Plackett model and goodness of approximation for Station India.

<table>
<thead>
<tr>
<th>$T_{02}$ marginal</th>
<th>$H_s$ marginal</th>
<th>$\psi^*$</th>
<th>$D^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-normal</td>
<td>Log-normal</td>
<td>6.3672</td>
<td>0.0281</td>
</tr>
<tr>
<td>Gamma</td>
<td>Log-normal</td>
<td>6.3672</td>
<td>0.0289</td>
</tr>
<tr>
<td>Log-normal</td>
<td>Weibull</td>
<td>6.5625</td>
<td>0.0306</td>
</tr>
<tr>
<td>Gamma</td>
<td>Weibull</td>
<td>6.6016</td>
<td>0.0312</td>
</tr>
<tr>
<td>Bivariate log-normal</td>
<td></td>
<td></td>
<td>0.0319</td>
</tr>
<tr>
<td>Log-normal</td>
<td>Gamma</td>
<td>6.4453</td>
<td>0.0340</td>
</tr>
<tr>
<td>Gamma</td>
<td>Gamma</td>
<td>6.5234</td>
<td>0.0348</td>
</tr>
<tr>
<td>Weibull</td>
<td>Log-normal</td>
<td>6.4854</td>
<td>0.0357</td>
</tr>
<tr>
<td>Weibull</td>
<td>Weibull</td>
<td>6.8359</td>
<td>0.0372</td>
</tr>
<tr>
<td>Weibull</td>
<td>Gamma</td>
<td>6.7188</td>
<td>0.0419</td>
</tr>
</tbody>
</table>

Fig. 4. Bivariate models for $(H_s, T_{02})$ data from Station India: (a) Plackett density with log-normal marginals; (b) Plackett density with log-normal $H_s$ marginal and Gamma $T_{02}$ marginal; (c) Plackett density with Weibull $H_s$ marginal and log-normal $T_{02}$ marginal; (d) bivariate log-normal density.
for the $H_S$-distribution appear also to give satisfactory fit to the data.

Keeping the marginal parameters fixed, then, we proceed to the ML estimation of the correlation parameter $\psi$. The ML estimates $\psi^*$ for each variant of the Plackett distribution are shown in Table 3. On the other hand, the Plackett minimum-variance estimate for $\psi$ (see eqn (27)) is $\psi = 5.33$, which is quite close to the ML estimates $\psi^*$ given in Table 3. In the same table also the measure $D^2$ of goodness-of-approximation is given for each Plackett variant. The presentation of the nine Plackett variants is in increasing order of the $D^2$ measure, i.e. in decreasing order of the goodness-of-approximation.

As expected, the Plackett distribution with log-normal marginals offers the best global fit ($D^2 = 0.0281$), better than the standard bivariate log-normal distribution ($D^2 = 0.0319$). The second best combination is that of Gamma for $T_{02}$ and log-normal for $H_S$, while the third best is the combination of log-normal for $T_{02}$ and Weibull for $H_S$. Figures 4(a), (b) and (c) show these three best Plackett distributions, while Fig. 4(d) shows the bivariate log-normal distribution. As is seen from this figure the Plackett variant with log-normal marginals provides a distribution quite similar, although not identical, to the bivariate log-normal distribution.

Finally, the global ML estimate $\theta^{**}$ has been calculated for the best-fit Plackett variant. The values of the $\theta^{**}$ components are shown in Table 4, in comparison with the previous estimates $\psi$ and $\theta^*$. It is seen that the difference between $\theta^*$ and $\theta^{**}$ is very small. This fact is shown in graphical form in Fig. 5, where the bivariate densities $f_{T_{02},H_S}^{(s)}(T_{02}, H_S; \theta^*)$ and $f_{T_{02},H_S}^{(s)}(T_{02}, H_S; \theta^{**})$ are shown. The solid contours correspond to $\theta^*$ while the dashed contours correspond to $\theta^{**}$, obtained by solving the global ML problem. The closeness of the contours illustrates the efficiency of the estimate $\theta^*$, obtained by maximizing separately the 'partial' log-likelihood functions $l_{T_{02}}^{(s)}(s_1; \theta_1)$, $l_{H_S}^{(s)}(s_2; \theta_2)$ and $l_{H_S}^{(s)}(s; \theta_1, \theta_2)$ (see Section 4).

6.2 Station Tromsøflaket

In a recent publication, Mathiesen and Bitner-Gregersen propose a conditional-distribution approach to the problem of analytic representation of $(H_S, T_{02})$ distribution. More specifically, they suggest a three-parameter Weibull distribution for $f_{H_S}$, combined with a two-parameter log-normal distribution for the conditional $f_{T_{02}|H_S}$. Furthermore, the two parameters of the $f_{T_{02}|H_S}$ distribution are represented as three-parameter model functions of $H_S$. Thus, finally, a nine-parameter model is obtained.

In the implementation of their model (called hereafter the M-B model) Mathiesen and Bitner-Gregersen concentrate on the accurate description of the steeper sea states, which may cause extreme structural responses. For the evaluation of their model they use data from three locations on the Norwegian continental shelf: Tromsøflaket, Utsira and Haltenbanken. Data from Station Tromsøflaket are used subsequently for a comparative study of the Plackett model and the M-B model.

Table 5 shows the $(H_S, T_{02})$ histogram of absolute frequencies of occurrence for Station Tromsøflaket. The data are based on a total of 15 605 wave measurements and they are also presented in the form of contour plots in Fig. 6.

The fit of the M-B model to Tromsøflaket data is shown in Fig. 7(a). Contours have been reproduced by using the parameter values provided by Mathiesen and Bitner-Gregersen. Although the particular model describes accurately the steeper sea states, it exhibits significant deviations for moderate and low wave heights. This behaviour is worse near the threshold of the
Table 5. (H_s, T_o2) histogram of absolute frequencies of occurrence for Station Tromsøflaket (After Mathiesen and Bitner-Gregersen)\(^\text{12}\)

<table>
<thead>
<tr>
<th>H_s (m)</th>
<th>0-0.5</th>
<th>0.5-1.0</th>
<th>1.0-1.5</th>
<th>1.5-2.0</th>
<th>2.0-2.5</th>
<th>2.5-3.0</th>
<th>3.0-3.5</th>
<th>3.5-4.0</th>
<th>4.0-4.5</th>
<th>4.5-5.0</th>
<th>5.0-5.5</th>
<th>5.5-6.0</th>
<th>6.0-6.5</th>
<th>6.5-7.0</th>
<th>7.0-7.5</th>
<th>7.5-8.0</th>
<th>8.0-8.5</th>
<th>8.5-9.0</th>
<th>9.0-9.5</th>
<th>9.5-10.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>T_o2 (s)</td>
<td>4</td>
<td>337</td>
<td>161</td>
<td>12</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6. ML estimates \(\delta_1, \delta_2\) of marginal parameters for Station Tromsøflaket

<table>
<thead>
<tr>
<th>Univariate model</th>
<th>(T_o2)-marginal</th>
<th>(H_s)-marginal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\alpha_1)</td>
<td>(\beta_1^*)</td>
</tr>
<tr>
<td>Log-normal</td>
<td>-0.8088</td>
<td>0.2115</td>
</tr>
<tr>
<td>Gamma</td>
<td>22.4920</td>
<td>0.0202</td>
</tr>
<tr>
<td>Weibull(^a)</td>
<td>29.5750</td>
<td>2.3926</td>
</tr>
</tbody>
</table>

\(\gamma_{T_o2} = 3.25\text{ s}; \gamma_{H_s} = 0.822\text{ m}.)

The estimated parameters \((\alpha_1, \beta_1^*)\) and \((\alpha_2, \beta_2^*)\) for the \(T_o2\) and \(H_s\) marginals, respectively, are given in Table 6, for each of the three univariate marginal models considered. The ML estimate \(\psi^*\) for each Plackett variant, and the corresponding measure \(D^2\) of goodness-of-approximation are given in Table 7. The presentation is again in increasing order of the \(D^2\) measure, i.e. in decreasing order of the goodness of approximation. The Plackett minimum-variance estimate for the correlation parameter is \(\psi = 10.4\), which is again quite close to the ML estimates given in Table 7.

Table 7. ML estimates \(\psi^*\) of the correlation parameter of the Plackett model and goodness of approximation for Station Tromsøflaket

<table>
<thead>
<tr>
<th>(T_o2) marginal</th>
<th>(H_s) marginal</th>
<th>(\psi^*)</th>
<th>(D^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-normal</td>
<td>Gamma</td>
<td>11.4840</td>
<td>0.0524</td>
</tr>
<tr>
<td>Gamma</td>
<td>Gamma</td>
<td>11.7190</td>
<td>0.0563</td>
</tr>
<tr>
<td>Bivariate log-normal</td>
<td></td>
<td>11.9730</td>
<td>0.0573</td>
</tr>
<tr>
<td>Log-normal</td>
<td>Log-normal</td>
<td>10.9380</td>
<td>0.0590</td>
</tr>
<tr>
<td>Log-normal</td>
<td>Weibull</td>
<td>11.7190</td>
<td>0.0724</td>
</tr>
<tr>
<td>M-B model</td>
<td></td>
<td></td>
<td>0.0812</td>
</tr>
</tbody>
</table>

Weibull marginal distribution which was taken as \(\gamma_{H_s} = 0.987\text{ m}.

The best-fit variant of the Plackett system, is the one with a Gamma \(H_s\) marginal and a log-normal \(T_o2\) marginal, and it is shown in Fig. 7(b). In Fig. 7(c) the Plackett variant with Weibull and log-normal marginals for \(H_s\) and \(T_o\), respectively, is shown, while Fig. 7(d) refers to the fit of the standard bivariate log-normal distribution\(^12\).
The M-B model describes steeper sea states better than does the Plackett best-fit variant. On the other hand, the Plackett model gives a better description of the data on a global basis; this is also illustrated in Fig. 8 where the marginal fit of the two models is presented. As can be seen in Fig. 8(a), the best-fit Plackett variant (log-normal for $T_{02}$) gives a better global fit of the $T_{02}$ marginal histogram than the M-B model. On the other hand, in Fig. 8(b) the M-B model fails to fit the $H_S$ data, for moderate values of $H_S$, whereas the best-fit Plackett variant (Gamma for $H_S$) describes the $H_S$ marginal histogram satisfactorily.

7 DISCUSSION AND CONCLUSIONS

In applied probabilistic and statistical work involving multivariate random variables, multivariate probability models play an important role. Such models, if suitably designed, can enhance statistical estimation, and serve for statistical prediction of derived quantities, simulation studies, etc.

In general, the construction of a multivariate probability model (other than the normal one and its translation system) is a difficult problem that can be treated by a variety of different methods. The suitability of one
method or another is dependent, to a large extent, on the kind of additional structure that we want to build in our model, which, in turn, is dependent on the specific information (analytic or empirical, quantitative or qualitative) that we have at our disposal. In the case where the model is to be fitted to empirical data, without any underlying theoretical background giving some evidence as regards its form, the most reliable information concerns the univariate marginals. Accordingly, it is very important that the multivariate model permits a satisfactory and directly controlled fit to all univariate marginal data. If this is the ease, the class of multivariate probability models with prespecified marginals (the Fréchet class) seems to be the most powerful and versatile system of candidate models.

A principal aim of this work was to bring into light such a philosophy for constructing multivariate probability models. Although our analysis has been restricted to the bivariate case, almost all the representations and techniques discussed here can be extended to the multivariate case as well. The basic advantages of probability models from the Fréchet class are:

- All univariate marginal models can be chosen in advance, in conformity with the marginal data, and the estimation of their parameters is made separately, on the basis of the corresponding marginal data set.
- As a consequence, additional univariate marginal data, that is, observations of only one component of the bivariate (multivariate) vector under study, can be taken into account in deciding the form and estimating the parameters of the marginal models from the marginal data.
- The dependence structure is controlled by an appropriate set of parameters, which are estimated from the multivariate data, after having determined the parameters of the marginal models.
- Various general alternative representations, as well as special bivariate and multivariate models, are available.

As an example, a special bivariate model, first constructed by Plackett,23 has been thoroughly studied and applied to the representation of the joint probability distribution of significant wave height and zero-upcrossing period. In comparison with the standard bivariate log-normal distribution, the Plackett model seems to offer a better performance regarding the description of the initial data, and an increased flexibility regarding the shape of the distribution, at the same computational cost, since the number of involved parameters remains unchanged (five parameters). In comparison with the M–B model, the Plackett model offers a better global description of the initial data. This fact is particularly seen in the marginals. On the other hand, the M–B model fits steeper sea states better than does the Plackett best-fit variant; this is due to the fact that Mathiesen and Bitner-Gregersen,12 in estimating the parameters, focus their attention on steep sea states, while, on the other hand, all the data were considered in the present case, in order to attain global fit to the data. Let it be added that, since the proposed model can incorporate any type of marginal distribution, it offers an increased flexibility for exploiting new univariate models to the bivariate case.
REFERENCES


Bivariate distributions with given marginals and wave climate


