

Probabilistic downtime estimation for sequential marine operations

Bruijn, Willem E.L.; Rip, Jolien; Hendriks, Antoon J.H.; van Gelder, Pieter H.A.J.M.; Jonkman, Sebastiaan N.

DOI

[10.1016/j.apor.2019.02.014](https://doi.org/10.1016/j.apor.2019.02.014)

Publication date

2019

Document Version

Final published version

Published in

Applied Ocean Research

Citation (APA)

Bruijn, W. E. L., Rip, J., Hendriks, A. J. H., van Gelder, P. H. A. J. M., & Jonkman, S. N. (2019). Probabilistic downtime estimation for sequential marine operations. *Applied Ocean Research*, 86, 257-267. <https://doi.org/10.1016/j.apor.2019.02.014>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' – Taverne project

<https://www.openaccess.nl/en/you-share-we-take-care>

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.



Probabilistic downtime estimation for sequential marine operations

Willem E.L. Bruijn^{a,b,*}, Jolien Rip^{a,b}, Antoon J.H. Hendriks^a, Pieter H.A.J.M. van Gelder^c, Sebastiaan N. Jonkman^b

^a Royal Boskalis Westminster N.V., Papendrecht, the Netherlands

^b Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Hydraulic Engineering, Delft, the Netherlands

^c Delft University of Technology, Faculty of Technology, Policy and Management, Safety and Security Science Group, Delft, the Netherlands

ARTICLE INFO

Keywords:

Simulation
Markov theory
Marine project operations
Offshore
Workability
Sequences
Linked
Chains
Probability
Synthetic
Time series
Downtime

ABSTRACT

A marine project consists of series of operations, with each operation subject to a predefined operational limit and duration. If actual weather conditions exceed the operational limit, the operation cannot be executed and hence downtime occurs. An accurate assessment of uncertainties and the expected downtime during a marine project is important in the tender and execution phase. This paper proposes a new probabilistic model for downtime estimation. It utilizes linked Markov chains that use actual metocean conditions to produce binary workability sequences for each operation. Synthetic time-series can be generated based on the statistics of the past observations and more project simulations are realizable, reducing the simulation uncertainty. The capabilities and limitations of the proposed approach are illustrated in a case study for a hypothetical project in the Tasman Sea.

1. Introduction

Nowadays, there is an increased interest in marine projects, such as the installation of offshore wind farms and platforms. These are often large-scale projects with high costs and relatively high levels of uncertainties associated with offshore weather conditions and marine operations. It is therefore important to improve the understanding and characterization of the uncertainties associated with the execution of marine projects.

A marine project typically consists of a series of operations. For instance, the installation of an offshore wind turbine foundation involves: sailing to project site, installing a monopile, placing transition piece and sailing back to harbor. Each operation is subject to a predefined operational limit and net duration, depending on the equipment being used. Operational limits can typically be related to threshold levels of environmental parameters such as wave height, wind speed and surface current or relevant combinations. If weather conditions exceed the operational limit, then the operation cannot be executed and hence downtime occurs. The net duration of an operation is defined as the time required to complete an operation without any delay. This is characterized by means of a deterministic value that optionally includes a safety margin for contingencies. Marine

contractors try to accurately estimate the uncertainties in the occurring conditions to obtain the expected downtime. This will be important for project management during the execution and to estimate the expected project costs (and the associated bandwidth) in the tender phase.

Currently, various approaches are being used to determine downtime for marine operations or marine projects. These are summarized in Table 1 and briefly described below. The most frequently used method to determine the workability is the wave scatter approach [1]. Workability is defined by the proportion of time that a time series is in the operable state, i.e. an operation can be executed. However, the wave scatter approach lacks information regarding the persistency. Persistence is the duration that one or more metocean parameters remain below or above a certain threshold [2]. The threshold is determined by the operation with its operational weather limit. Models including persistency are developed by [3–5] with empirical distribution fits, by [6] with the Markov theory and by [7] using the Equivalent Storm Model approach.

However, none of these methods allow to analyze the downtime for a series of operations, or to incorporate more than two metocean parameters in a project. The downtime of series of operations can be either determined analytically by means of a probabilistic network and evaluation technique [11] or an event tree [12]. However, analytical

* Corresponding author.

E-mail address: welbruijn@gmail.com (W.E.L. Bruijn).

<https://doi.org/10.1016/j.apor.2019.02.014>

Received 19 August 2018; Received in revised form 27 January 2019; Accepted 20 February 2019

Available online 13 March 2019

0141-1187/ © 2019 Published by Elsevier Ltd.

Table 1
Overview of existing approaches for downtime analysis in marine projects [26].

Method	Pros	Cons	Literature
Analytical			
Joint probability distribution / wave scatter	Simple, fast and gives a quick overview of the operational limits relative to the prevailing weather conditions.	Not able to take weather windows or pattern of transitions between states into account.	[1]
Empirical persistency distribution	Lot of literature for H_s application.	Not (yet) applied for > 2 metocean parameters or multiple operations, long time series required.	[2,4,8,3,9]
Markov chain theory	Shorter time series required.	Not (yet) applied for > 2 metocean parameters or multiple operations.	[6,10]
Equivalent Storm Model	Simple expressions for downtime and missed energy	Limited to H_s only	[7]
Probabilistic network	Analytical downtime distribution for a marine project.	Duration per path assumed to be normal distributed and operation duration assumed to be independent (not always true).	[6,11]
Event / decision tree	Closed-form expression for downtime.	Trees might become too large for marine projects.	[12]
Simulation			
No stochastic model (direct use of hindcast data)	Easy (no model construction), no parametric uncertainty and relatively low model uncertainty.	Confidence bounds on downtime distribution large due to limited sample size (high simulation uncertainty)	[1,13,14,15,16]
Stochastic models:			
• Resampling	No parameter estimation required.	Persistency patterns can appear.	[16]
• Multivariate distribution	Exact values of metocean parameters modelled.	High parametric and model uncertainty for 2+ metocean parameters.	[17,18]
• Markov (discretized metocean parameter values as states)	CDF of metocean parameters and persistence statistics successfully reproduced.	High parametric uncertainty for 2+ metocean parameters.	[19,20]
• Markov (workability states)	Any combination of metocean parameters possible, relatively low simulation uncertainty.	Relatively higher parametric and model uncertainty.	[6], this research for multiple sequential operations.

analysis becomes complex in case of a sequence of operations and/or a lot of metocean parameters dictating the operational limits. It can also be determined by means of a simulation/sampling study [1] [13]. The main downside of a simulation is that sufficient metocean data is required for multiple simulations of a project to reach a reliable solution. The associated uncertainty is called the ‘simulation uncertainty’ [21]: the uncertainty related to the number of replications. Simulation uncertainty can be decreased by increasing the number of realizations.

A simulation study can be performed with real (observed) data or synthetic data produced by a stochastic model. Since a project is bound to a start date and weather conditions generally vary over the year, the number of simulations on observed data is proportional to the number of years recorded in the data. Metocean datasets are being recorded or hindcasted for approximately 15–35 years [22], yielding in only 15–35 project simulations. This generally results in a high simulation uncertainty in the current approaches for downtime estimation, and thus an uncertain project cost estimation. As far as known by the authors, current literature does not provide a stochastic model to estimate downtime of marine projects, consisting of multiple operations which are defined by its operational limits and durations. Monbet et al. [23] conducted a survey of stochastic models for wind and sea state time series. Most models only allow one or two metocean parameters, while more parameters may influence the operational limit in practice. Hence, these stochastic models are less applicable for downtime estimation.

To overcome these limitations, this paper introduces a new method in which synthetic project simulations are generated with Monte-Carlo Markov Chains based on the statistics of the hindcasted data. Within this approach there is no limit for the number of metocean parameters considered. With a synthetic model more project simulations are realizable, which reduces the simulation uncertainty, i.e. the uncertainty associated with the number of repetitions.

Next to the simulation uncertainty, two other important uncertainties are present: ‘parametric’ and ‘model’ uncertainty. The parametric uncertainty is the uncertainty related to the estimation of the input parameter(s) of the stochastic model due to the limited data

sample and parameter estimation method [24,21,25]. The lower the number of input parameters and/or the more data on which the parameter estimation is based, the lower the parametric uncertainty. The model uncertainty is related to non-statistical errors due to abstraction [24,25]. The more realistic the model, the lower the model uncertainty. This paper proposes only a quantification for the simulation uncertainty. In the discussion (Section 3.3) the impact of the other uncertainties will be qualitatively elaborated. The probabilistic model framework with linked Markov chains as presented in this paper was developed in the research report [26] and is further developed by Bruijn [27].

The remainder of the paper is structured as follows. In Section 2 the model framework is introduced. The ability of the model to generate synthetic project simulations from which the downtime can be determined, is explained. The model is applied with a hypothetical project and a metocean hindcasted dataset from the Tasman Sea in Section 3. The discussion of this research is presented in Section 4. Finally, Section 5 concludes with conclusions and recommendations.

2. Model overview

2.1. General

Markov chains have been used in various fields, such as the modelling of queuing systems, DNA sequences, financial risk, and many other practical applications [28]. Markov chains may also be used to model sea states [6,19,20]. In this section the model will be explained which is able to produce synthetic marine project simulations with Markov chains. The breakdown structure in Fig. A1 (Appendix A) shows the process of the model developed in this research.

The model abstracts the actual metocean conditions into workable states ‘1’ and non-workable states ‘0’ depending on the operational limit (see Fig. 1). So-called a binary ‘workability-array’ is created from the hindcasted data and the operational limit(s). Hence, the information about the actual metocean parameter is lost. Other delay factors apart from weather conditions are not considered in this paper. Based on the

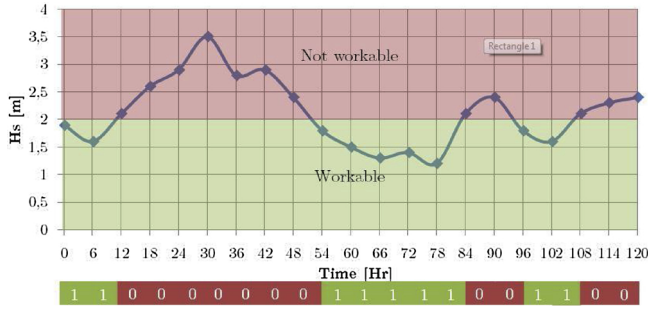


Fig. 1. An example of how to obtain the workability-array from the significant wave height time series and an operational limit ($H_s \leq 2\text{ m}$) [26].

statistics of the hindcasted 2-state workability-array, the model is able to produce binary ‘workability sequences’ stochastically for each operation. Also when an operation is limited to two or more metocean parameters, the model can still produce binary time-series for the specific operation without modelling dependencies between the parameters.

From the hindcasted workability-array Markov transition probabilities between the states can be estimated with the maximum likelihood estimate (MLE) [29] in Eq. (1).

$$\hat{P}_{ij} = \frac{N_{ij}}{N_i^*}, \quad \forall i, j \in S \quad (1)$$

Where N_{ij} is the number of observed transitions from state i to j and N_i^* is the number of transitions starting from state i ($N_i^* = \sum_j N_{ij}$, $\forall i, j \in S$). The limiting probability π_j is the probability that a process will be in state j after a large number transitions, and it can be computed with Eq. (2) [30].

$$\pi_0 = \frac{P_{10}}{P_{10} + P_{01}}, \quad \pi_1 = \frac{P_{01}}{P_{10} + P_{01}} \quad (2)$$

The workability sequence consist of time steps $t = 0, 1, \dots, T$, where discrete random variables $X_0, X_1, \dots, X_t$ represent the ‘state’ or ‘workability’ of the process at each time step t . If the process is in state 0 at time step t (i.e. $X_t = 0$), the operational limit is violated and the operation cannot be executed. Vice versa, a state of 1 is assigned if the operation can be executed. The proposed model is therefore a discrete time, discrete state space Markov chain. The following equation holds if a future state depends on the past u states in a Markov chain [30]:

$$\begin{aligned} P(X_{t+1} = j | X_t = i, X_{t-1} = i_{t-1}, \dots, X_0 = i_0) \\ = P(X_{t+1} = j | X_t = i, \dots, X_{t+1-u} = i_{t+1-u}) = P_{ij}(t), \quad \forall i, j \in S, \quad t \\ = 0, 1, \dots, T \end{aligned} \quad (3)$$

2.2. Assumptions

The following requirements and assumptions have been set:

Requirements: The synthetic dataset should show the same characteristics as the original data:

- Seasonality is respected.
- Persistency of sequential workable time steps is respected.
- Overall workability is respected.
- Dependencies between operations are respected.

Assumptions

- There are no long-term trends in the data (assumption of no climate change), only seasonal inhomogeneity is considered.
- The operational limit is strict: a time step is either workable or not.
- The net duration of an operation is deterministic.

- The statistics within the original dataset are assumed to represent reality.
- Only 1 metocean dataset is used, thus sailing hours are projected only for this location.

2.3. Metocean data and project planning

A project planning consists of a sequence of operations, where each operation is defined by its operational limit and net duration. Also, the relationship between operations are included, as some operations must start directly after the preceding operation is completed and others do not.

Hindcasted datasets of metocean parameters near project locations are used as input data for the model. These datasets consist of approximately 15–35 years of data, with generally 1-hour, 3-hour or 6-hour intervals. With the dataset and the operational limits the binary workability sequences are created. The probabilities can be calculated more accurately when more data is available. Therefore, one should always strive to use the largest available dataset near the project location.

2.4. Seasonality

Metocean conditions are subject to seasonality, hence two approaches are introduced to incorporate these. The first approach is called ‘piecewise time-homogeneous’ and the second approach is called ‘non time-homogeneous’.

In the piecewise time-homogeneous approach the year is divided into periods (e.g. seasons or months) and the transition probabilities are estimated separately for these periods with equation (1) and are assumed to be constant over the periods.

In the non-time-homogeneous approach the transition probabilities vary over the year depending on the day of the year. A discrete non-parametric kernel estimator is applied to estimate the $P_{ij}(t)$ [31], which gives generally more weight to transitions near the calendar day of interest and the ones further away lower weightage. The resulting kernel estimators for transition probabilities $\hat{P}_{01}(t)$ and $\hat{P}_{10}(t)$ are given:

$$\hat{P}_{01}(t) = \frac{\sum_{k=1}^{N_{01}^*} K\left(\frac{t-t_{01k}}{h_{01}}\right)}{\sum_{k=1}^{N_0^*} K\left(\frac{t-t_{0k}}{h_{01}}\right)} \quad \hat{P}_{10}(t) = \frac{\sum_{k=1}^{N_{10}} K\left(\frac{t-t_{10k}}{h_{10}}\right)}{\sum_{k=1}^{N_1^*} K\left(\frac{t-t_{1k}}{h_{10}}\right)} \quad (4)$$

Where in Eq. (4) N_{01} = the number of transitions from state 0 to 1 in the workability-array; N_0^* = the number of transitions starting from state 0; $K(\cdot)$ = the kernel function (Eq. (5)); h_{01} = a kernel bandwidth; t = a calendar day; t_{01k} = the day indices of transitions from 0 to 1 in the data; t_{0k} = the day indices of transitions starting from 0. The estimates on calendar day t are hence obtained by using the information of the days in the range of $[t - h, t + h]$. Since the Markov chain is ergodic the other probabilities can be obtained by $\hat{P}_{00} = 1 - \hat{P}_{01}$ and $\hat{P}_{11} = 1 - \hat{P}_{10}$. The discrete kernel function is given by Rajagopalan et al. [31]:

$$K(x) = \frac{3h}{1 - 4h^2} (1 - x^2), \quad \text{for } |x| \leq 1 \quad (5)$$

By means of a least squared cross validation procedure the kernel bandwidths are determined:

$$LSCV(h_{01}) = \frac{1}{N_0^*} \sum_{i=1}^{N_{01}} (P_{01,init}(t_i) - \hat{P}_{01,-t_i}(t_i))^2 \quad (6)$$

$$LSCV(h_{10}) = \frac{1}{N_1^*} \sum_{i=1}^{N_{10}} (P_{10,init}(t_i) - \hat{P}_{10,-t_i}(t_i))^2 \quad (7)$$

Where $\hat{P}_{01,-t_i}(t_i)$ = the estimate of the transition probability on day t_i , dropping the information on day t_i . $P_{01,init}(t_i) = \frac{N_{01}(t_i)}{N_0^*(t_i)}$, where $N_{01}(t_i)$ and $N_0^*(t_i)$ are the number of transitions from 0 to 1 on day t_i and the number of transitions starting from 0 on day index t_i respectively. The

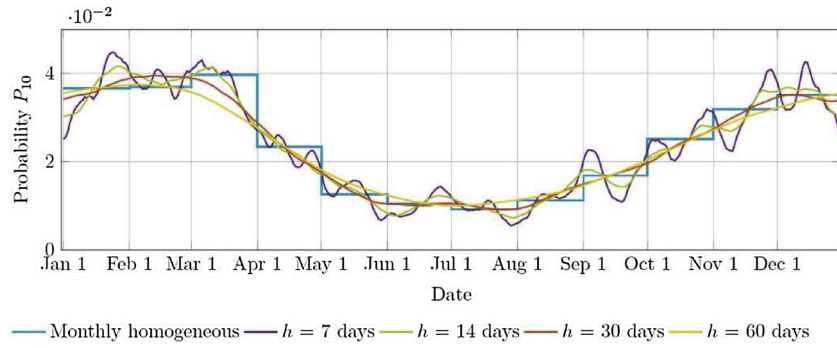


Fig. 2. Comparison between first order monthly constant transition probabilities (blue) and the non-time-homogeneous transition probabilities with different kernel bandwidths. NOAA dataset with an operational limit of $H_s \leq 2.5 \text{ m}$ & $U \leq 10 \text{ m/s}$ [26].

bandwidth is investigated for periods from 7 to 30 days. It is considered that for a smaller bandwidth too little data is used to make a justified estimate, and for a larger bandwidth the estimates become too ‘smoothed’ as can be seen in Fig. 2 with bandwidth of 60 days. In Fig. 2 the transition probability P_{10} is presented over the year calculated with both seasonality approaches and with different bandwidths is given. This is done based on a measured dataset by National Oceanic and Atmospheric Administration (NOAA) over the period 1990–2011 with a temporal resolution of 1 h at a longitude of 78.493 °W and a latitude of 28.872 °N.

In order to determine which seasonality approach most accurately describes the hindcasted data, the homogeneity test is performed. The transition probabilities per week within a month should be more or less the same, if the assumption of monthly piecewise stationarity is correct [32]. The monthly workability sequences are subdivided into $Y = 4$ different sub-intervals (weeks per month). The transition probabilities per week (sub-interval) are tested with the transition probabilities per month. The following hypotheses are defined:

- $H_0: P_{ij}(y) = P_{ij}(m) \quad \forall i, j \in S, y = 1, 2, 3, 4, m = 1, 2, \dots, 12$
- $H_1: P_{ij}(y) \neq P_{ij}(m) \quad \forall i, j \in S, y = 1, 2, 3, 4, m = 1, 2, \dots, 12$

Where $P_{ij}(y)$ denotes the transition probability from state i at time t to state j at time $t + 1$ during sub-interval y (week). This transition probability holds for $[t_y, t_y + \Delta y]$, where t_y is defined as the first time step in sub-interval y with length Δy . $P_{ij}(m)$ denotes the transition probability from state i at time t to state j at time $t + 1$ during interval m (month). This transition probability holds for $[t_m, t_m + \Delta m]$, where t_m is defined as the first time step in interval m with length Δm . The transition probabilities during sub-interval y or interval m are calculated with the maximum likelihood estimate from Eq. (1).

The Chi-square test is used to test the null-hypothesis [29]:

$$X_i^2 = \sum_{y=1}^Y \sum_{j \in S} \frac{N_{ij}(y)(\hat{P}_{ij}(y) - \hat{P}_{ij}(m))^2}{\hat{P}_{ij}(m)}, \quad \forall i, j \in S \quad (9)$$

The limiting X_i^2 distribution has $(n_s^u - 1) \cdot Y$ degrees of freedom (u is the order of the Markov chain and n_s is defined as the number of states). Summing over all X_i^2 , the total test statistic X_t^2 has a limiting X_t^2 distribution with $(n_s^u - 1) \cdot Y \cdot n_s^u$ degrees of freedom. A small number (10^{-10}) is added to the number of transitions for smoothing to avoid $\hat{P}_{ij}(y) = 0$. If the null hypothesis is true, the hindcasted data is considered to be piecewise time-homogeneous. Alternatively, the non-time-homogeneity is assigned to the model.

2.5. Time-dependency

All metocean conditions are subject to autocorrelation, which can be captured in the time-dependency of a Markov chain. In a first-order

Markov chain the value of X_{t+1} only depends on the value of X_t . By using a higher Markov chain order the value of X_{t+1} depends on the value of $X_t, X_{t-1}, \dots, X_{t-(u-1)}$ with u denoting the order. The maximum likelihood estimate can be obtained by:

$$\hat{P}_{ij \dots kl} = \frac{N_{ij \dots kl}}{N_{ij \dots k}}, \quad \forall i, \dots, k \in S \quad (9)$$

Where $N_{ij \dots kl}$ denotes the number of observed transitions from state $ij \dots k$ to state l and $N_{ij \dots k}^*$ denotes the number of observed transitions that start from state $ij \dots k$ ($N_{ij \dots k}^* = \sum_j N_{ij \dots k}, \quad \forall i, j, k \in S$). The additional history of a higher order can make the predictions more accurate, but on the other hand more parameters need to be estimated which grows exponentially with $(n_s - 1)n_s^u$. Where n_s = the number of states and u = the Markov chain order. Additionally, the limiting probabilities are calculated differently for higher orders:

$$\hat{\pi}_{jk \dots l} = \frac{1}{T - u} \sum_{t=1}^{T-u} 1_{\{X_t=j, X_{t+1}=k, \dots, X_{t+u-1}=l\}} \quad (10)$$

where

$$1_{\{X_t=j, X_{t+1}=k, \dots, X_{t+u-1}=l\}} = \begin{cases} 1 & \text{if } X_t = j, X_{t+1} = k, \dots, X_{t+u-1} = l \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

To show how the Markov chain order influences the persistency, a North sea dataset is used. This dataset is collected by Boskalis at (53.91 °N, 2.15 °E), consisting of 23 years (from 1992 till 2015) sampled with an interval of 3 h. In Fig. 3 a cumulative distribution function (CDF) of the persistency is displayed of the North sea dataset with an operational limit of $H_s \leq 2 \text{ m}$. The most upper line (blue) shows the cumulative persistency distribution of the hindcasted data. It can be noticed that the cumulative persistency distribution of the synthetic

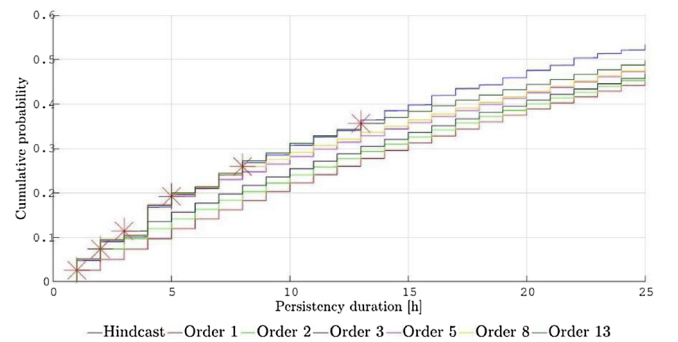


Fig. 3. Increasing the Markov chain order the cumulative distribution of persistency of H_s remaining below 2 m at the North sea location converges to the hindcasted data. The red markers indicate the splitting points where the generated dataset splits off the hindcasted dataset [27].

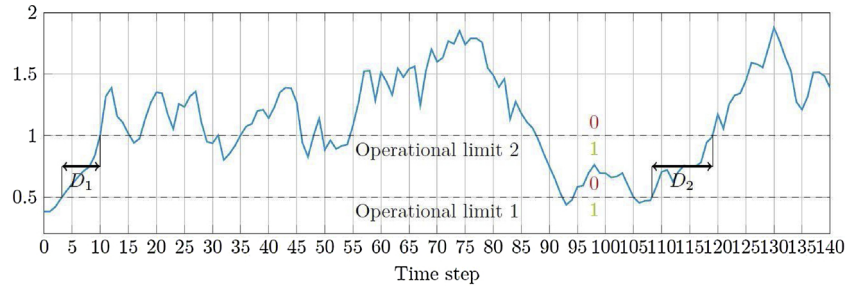


Fig. 4. Example how the workable influence periods for an operational limit $H_s \leq 1\text{ m}$ following operational limit $H_s \leq 0.5\text{ m}$ are determined [26].

datasets converges to the hindcasted data by applying a higher Markov chain order. In other words, it can be stated that a higher Markov chain order better preserves the persistency distribution. The red markers indicate the splitting points where the synthetic datasets split off the hindcasted data. It can be seen that the persistency distribution is preserved until the duration of the persistency (hours) is equal to the Markov chain order. From this point the curve of the synthetic dataset releases from the hindcasted data curve, and persistency is over-estimated. An even higher Markov chain order than 13 is needed to preserve a more accurate persistency distribution of the North sea dataset with operational limit of $H_s \leq 2\text{ m}$. This increases the risk that the model will tend to replicate the exact hindcasted data, leading to no new information and making the model of no added value. However, the model-user should analyze which order and which seasonality approach should be used per location and per project. It is recommended not to use a higher Markov chain order than the longest net duration (in hours) of all operations in the project, since longer persistency windows are not being used.

2.5.1. Order test

The order-test can be used to determine which Markov chain order describes the workability sequences most accurate for the piecewise time homogeneous approach. In this paper the non-time-homogeneous approach is given with a first-order Markov chain, therefore this test will only be performed if the homogeneity test is accepted. Transition probabilities for Markov chain order u and Markov chain order $u+1$ should be more or less the same, if the assumption of Markov chain order u is correct [29]. In other words, the last state before u states should not have a significant influence on the transition probabilities.

- H_0 : The Markov chain order is of order u , implying that $P_{ij...kl} = P_{j...kl}$, $\forall i, j \in S$ (the workability sequence $ij...k$ covers the last $u+1$ states and the workability sequence $j...k$ covers the last u states, l is the next state)
- H_1 : The Markov chain is not order u , hence $P_{ij...kl} \neq P_{j...kl}$

A Chi-squared test can be used to test the null hypothesis, which is defined as:

$$X^2_{j...k} = \sum_{i \in S} \sum_{l \in S} \frac{N_{ij...kl}^* (\hat{P}_{ij...kl} - \hat{P}_{j...kl})^2}{\hat{P}_{j...kl}}, \quad \forall i, j, \dots, k \in S \quad (12)$$

Where $\hat{P}_{ij...kl}$ and $\hat{P}_{j...kl}$ are calculated with Eq. (10) using order u and order $u-1$ respectively. The limiting X^2 distribution has $(n_s - 1)^2$ degrees of freedom. Summing over all X^2 , the total test statistic X^2 has a limiting X^2 distribution with $(n_s - 1)^2 \cdot n_s^u$ degrees of freedom. A small number (10^{-10}) is added to the number of transitions for smoothing to avoid $\hat{P}_{ij} = 0$. If the null-hypothesis is true, then the workability sequence has a Markov chain order u .

2.6. Linking Markov chains

With the previously described theory the model is able to produce time-series for single operations. However, marine operations consist of multiple sequential operations and therefore the Markov chains of these individual operations have to be linked. For that purpose two concepts are introduced: the ‘influence period’ and the ‘cross-transition probability’.

2.6.1. Influence period

The general idea of the influence periods is to ensure that the speed at which a sea regime changes in the synthetic data is physically feasible. For example, a workable time step t in an operation with operational limit $H_s \leq 0.5\text{ m}$ followed by a non-workable time step $t+1$ in an operation with limit $H_s \leq 3\text{ m}$ implies a change from $H_s \leq 0.5\text{ m}$ to $H_s > 3\text{ m}$ in 1 h, which is very unlikely. Hence, a minimum duration of required workable time steps need to be implemented in the second operation q after the completion of the first operation p , this duration is called the workable influence period $D^1(p, q)$. Vice versa, a change from $H_s > 3\text{ m}$ to $H_s \leq 0.5\text{ m}$ in 1 h is unlikely as well, hence the non-workable influence period $D^0(p, q)$ is introduced. This is the minimum required non-workable time steps that need to be implemented in the second operation q after the crossing of non-workable time step to a workable time step in operation p .

The workable influence period is determined by collecting the duration where operation p crosses from a workable time step to a non-workable time step, till operation q crosses from a workable time step to a non-workable time step. From this collection the influence periods corresponding to an exceedance probability of 0.98 are implemented, as it is assumed that no more than 2% in the data contains errors and extreme events are avoided. The workable influence periods are implemented for all succeeding operations after the completion of operation p .

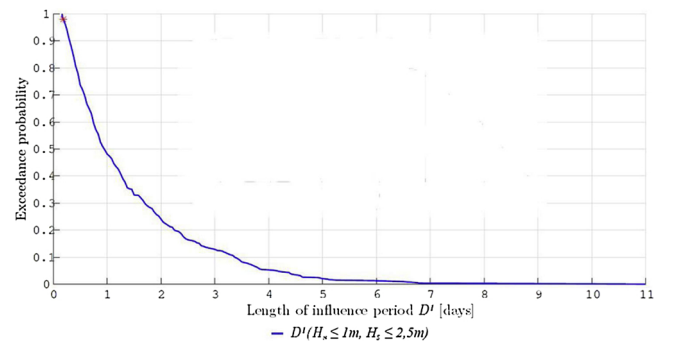


Fig. 5. Exceedance probability of the workable influence period $D^1(H_s \leq 1\text{ m}, H_s \leq 2.5\text{ m})$ with a red marker at an exceedance probability of 0.98 corresponding with an influence period of approximately 5 h [27].

Unlike the workable influence period, the non-workable influence period is scheduled for each operation that follows after a 0/1 crossing in the workability sequence of operation p . Similarly to the workable influence period, the collection of non-workable influence periods happens in the same way and non-workable influence periods corresponding with an exceedance probability of 0.98 are used when operation p crosses from a non-workable to a workable time step.

In Fig. 4 an example is given for the D^I values for an operation with a limit $H_s \leq 0.5 m$, followed by an operation with a limit $H_s \leq 1 m$. Such time periods from passing the 1/0 boundary of operation p , to the passing of the 1/0 boundary of operation q are collected in an empirical exceedance probability curve. Fig. 5 presents the empirical exceedance probability curve of the workable influence periods determined by operations p and q with operational limits of $H_s \leq 1 m$ and $H_s \leq 2.5 m$ respectively. This not only applies for the first following operation, but for all the following operations an exceedance probability curve is determined. After operation p is completed the influence periods corresponding with an exceedance probability equal to 0.98 are scheduled for each operation that follows. After the implementation of the workable influence periods, the Markov chain continues with the regular transition probabilities for the next operation.

2.6.2. Cross-transition probability

If no influence period is scheduled, the cross-transition probability $\hat{P}_{ij}^{pq}$ is used to capture the dependency between the operations. The cross-transition probability characterizes the likelihood that the next state in operation q is either workable or non-workable based on the previous state of operation p . Similarly as the regular transition probabilities, the cross-transition probabilities can be derived from the observed transitions as follows:

$$\hat{P}_{ij}^{pq} = \frac{N_{ipjq}}{N_{ip}^*}, \quad \forall i, j \in S \quad (13)$$

Where N_{ipjq} is the number of observed transitions from state i in the workability-array of operation p to state j in the workability-array of operation q and N_{ip}^* is the number of transitions that start from state i in operation p ($N_{ip}^* = \sum_{j \in S} N_{ipjq}$, $\forall i \in S$). This equation can be extended for higher orders as well and piecewise stationarity is assumed for this approach.

2.7. Coupled operations

When an operation *must* start *directly* after the preceding operation has completed, i.e. if no delay is allowed between two successive operations, only the preceding operation is called a coupled operation. In order to find a weather window which allows a coupled operation and the successive operations to be performed, the model generates the workability sequences for both operations simultaneously. Therefore, 4 states are defined as shown in Table 2, and their transition probabilities are derived with Eq. (1) with $S = \{0,1,2,3\}$. The same approach is applied when there are more sequentially coupled operations, resulting in a higher value for S .

Table 2

The workability states for a coupled operation and the succeeding operation.

Coupled operation	Subsequent operation	Modelled state value
0	0	0
0	1	1
1	0	2
1	1	3

2.8. Simulation uncertainty

The simulation uncertainty is determined by the number of project simulations: increasing the number of project simulations, the simulation uncertainty decreases. The outcome of the model is the distribution of the project duration or downtime duration. These can be derived by an empirical cumulative distribution function (ECDF), which is defined as:

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n 1\{X_i \leq x\}, \quad (14)$$

where

$$1\{X_i \leq x\} = \begin{cases} 1 & \text{if } X_i \leq x \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

In this equation n determines the number of project simulations. The simulation uncertainty is quantified by the Dvoretzky-Kiefer-Wolfowitz (DKW) inequality, which determines the confidence bounds for $\hat{F}_n$. With $\hat{F}_n$ the lower (L) and upper (U) confidence bounds are defined as [33]:

$$L(x) = \max \left\{ \hat{F}_n(x) - \sqrt{\frac{1}{2n} \ln \left(\frac{2}{\alpha} \right)}, 0 \right\} \quad (16)$$

$$U(x) = \min \left\{ \hat{F}_n(x) + \sqrt{\frac{1}{2n} \ln \left(\frac{2}{\alpha} \right)}, 1 \right\} \quad (17)$$

Then, for any CDF F and all n

$$P(L(x) \leq F(x) \leq U(x)) \geq 1 - \alpha \quad \forall x \in \mathbb{R} \quad (18)$$

Where $1 - \alpha$ is the probability at each point x that $F(x)$ lies within the confidence bounds. Fig. 6 shows that increasing the sample size n results in smaller maximum distances between $\hat{F}_n(x)$ and $L(x)$ and $U(x)$ in order of $O(\frac{1}{\sqrt{n}})$. Note that, the larger the sample size n develops, the less important the α -value gets as the lines converge. It is recommended to generate in the order of 1000 project simulations, because this reduces the probability that an outcome lies outside the confidence bounds to a value below of 5%, which is commonly accepted.

2.9. Validation

For validation purposes, the workability percentage and persistency distribution of the generated and hindcasted workability sequences can be compared. To check whether the workability is respected, the root mean squared error (RMSE) can be computed of the monthly generated workability sequences (indicated by $\hat{W}_m$) compared to the monthly workability of the hindcasted workability sequences (W_m):

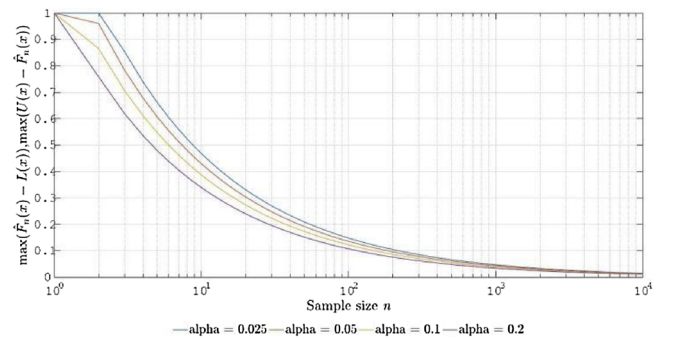


Fig. 6. Relationship of the maximum distance between $\hat{F}_n(x)$, lower bound $L(x)$ and upper bound $U(x)$ versus sample size n for multiple α -values [26].

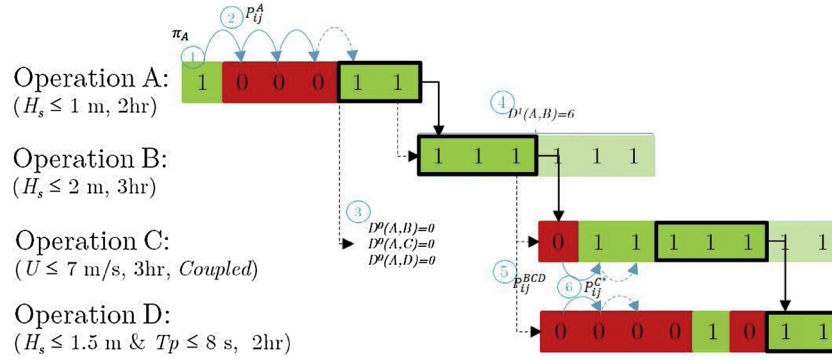


Fig. 7. Visual schematization of the model procedure of a hypothetical project consisting of 4 operations.

$$RMSE_{workability} = \sqrt{\frac{\sum_{m=1}^{12} (\hat{W}_m - W_m)^2}{12}} \quad (19)$$

To check whether the persistency of the hindcasted is respected in the generated workability sequences, the mean, standard deviation and the two-sample Kolmogorov-Smirnoff test can be computed for each operation. Suppose $L_1, \dots, L_n$ are the lengths of subsequent 1 s in the workability sequence. The ECDF is then computed by using Eq. (11), with X_i replaced by L_i . The two-sample Kolmogorov-Smirnoff test is as follows calculated [34]:

$$D_L = \max_x |\hat{F}_{L,obs}(x) - \hat{F}_{L,model}(x)| \quad (20)$$

2.10. Schematization

The model described in the previous sections has been implemented in MATLAB code. This allows a relatively short run time (order of 1 h) on a standard computer for a project consisting of 18 operations where 1000 stochastic project simulations are generated.

In Fig. 7 a schematization of a single hypothetical project simulation is depicted, where the blue numbers correspond with the numbered list below. In this small project 1st order Markov chain is applied.

- 1 The first symbol is produced with the limiting probabilities for operation A, as can be calculated with Eq. (2).
- 2 The workability sequence continues with the transition probabilities of operation A, which are calculated with Eq. (1), until the net duration of two hours is reached. The initiation and completion of the operation is framed with a black rectangle.
- 3 In case the crossing from a non-workable (0) to workable (1) state occurred in the workability sequence, the model will generate the non-workable influence periods $D^0(p, q)$ for the next operation. In this project there were no non-workable influence periods.
- 4 For all succeeding operations the influence period $D^1(p, q)$ of operation A is determined. Only operation B is influenced with 6 workable time steps due to its less strict operational limit. The other operations are not influenced at all, because of the different operational limits. The net duration of operation B fits within the scheduled influence period, and from this time step the project simulation continues.
- 5 For all succeeding operations the influence periods of operation B are determined, which are calculated to be zero. Since no influence period is scheduled for operation C and D and operation C is defined as a coupled operation (operation D has to start directly after completing operation C), the simulation continues with the cross-transition probability P_{ij}^{BCD} . The cross-transition probability is calculated with Eq. (10).

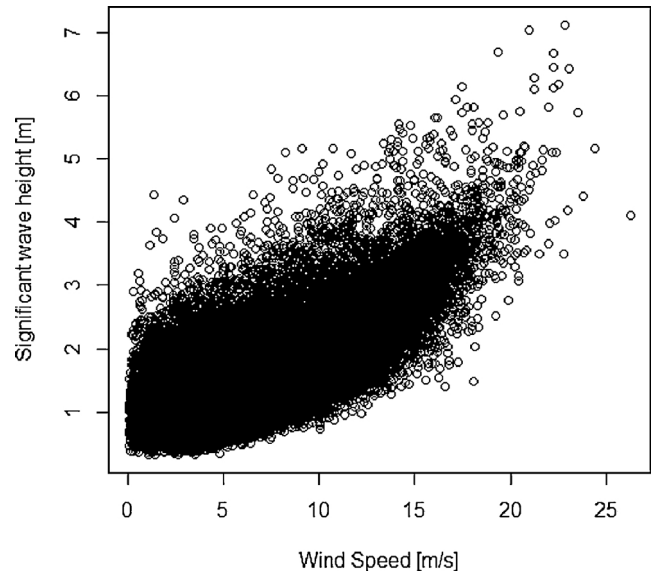


Fig. 8. Scatterplot of the significant wave height [m] and the wind speed [m/s] in the Tasman sea.

- 6 The workability sequence of operation C and D continue with the transition probabilities P_{ij}^C , which generates states simultaneously for operation C and operation D until the net duration for both operations is reached without downtime in between them. Note that operation C could not start any earlier because of the non-workable states in operation D.

These time steps summarize how one project is simulated, which resulted in a project duration of 17 time steps (hours) in this example. The downtime is calculated by subtracting the total net duration of the project duration which is $17 - 10 = 7$ time steps (hours).

3. Application and results

3.1. Model configuration

The model requires two input files, namely a metocean hindcasted dataset and a project planning. In this paper a metocean dataset from the Tasman Sea gathered at a longitude of 148°E and a latitude of 38.5°S is used. This dataset has been obtained by Boskalis and it

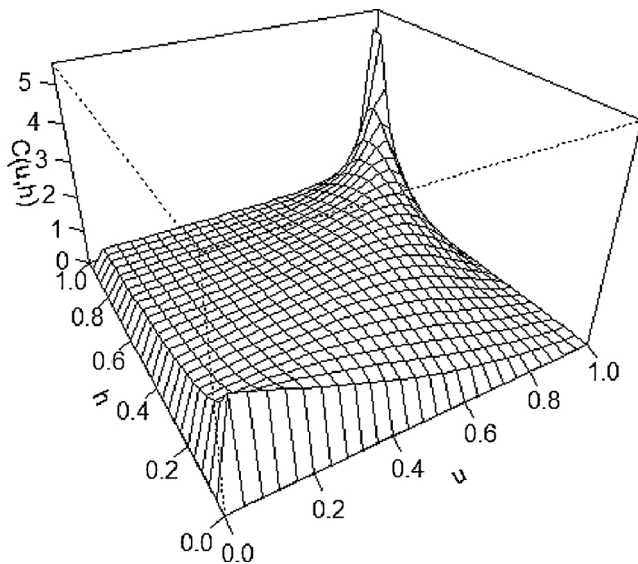


Fig. 9. Tawn T2 copula density plot of the significant wave height and wind speed in the Tasman sea.

consists of 24 years of data (1992 – till 2016) sampled with an interval of 3 h. Figs. 8 and 9 depict a scatterplot of the significant wave height (H_s) and wind speed (U), and the dependency between these variables with the Tawn T2 copula.

The other input file is the hypothetical project planning for an installation of a foundation of a wind turbine from Table 3. Operation 3b and 5 are not considered in the simulation, because operation 3b has no operational limit and operation 5 requires the water depth which is not measured.

The non-time-homogeneity approach with first-order Markov chain is applied for the simulation. January 1 is chosen to be the project start date.

3.2. Model performance

As explained in Section 2.8 the outcome of the model is the distribution of the project duration or downtime. Fig. 10 shows the downtime of the hypothetical project. The results indicated in red are determined with the model configurations of Section 3.1 and a sample size n (project simulations) of 1000. The results indicated in blue are determined by simulating the hypothetical project of Table 3 on the hindcasted data of 23 years, thus 23 project realizations. The

confidence bounds indicate the simulation uncertainty, calculated with Eq. (15). A further discussion of the results is given in the next section.

3.3. Discussion of results

With the new model the simulation uncertainty is significantly reduced, this is shown by the decrease of bandwidth in Fig. 6. On the other hand, parametric and model uncertainty are introduced which are not quantified. The model settings determine the degree of uncertainty, for example, the non-time homogeneity method results in a higher parametric uncertainty (because more transition probabilities need to be determined) and lower model uncertainty (due to more accurate estimations). The model and parametric uncertainty work in different directions, when one increases the other one decreases. For this simulation the non-time homogeneity approach is used, because it describes the probabilities more accurate than the piecewise time-homogeneous approach according to the homogeneity test. On the other hand, the non-time homogeneous method works with a 1st-order Markov chain which overestimates the persistency distribution, making it less accurate. Also, it is noted from Fig. 10 that there are substantial differences in expected outcomes between the two models, especially for downtime values between 3 and 12 days. This shows how the choice of the modelling concept can influence outcomes. Also, the original model (based on 23 years of data) shows very large bandwidth in project time for a given probability value (e.g. between 6 and 21 days for an 80th percentile value), making probabilistic planning difficult. The new model is in that respect more suitable for probabilistic project planning.

There were no coupled operations scheduled for this hypothetical project. The parametric uncertainty increases significantly when (sequentially) coupled operations are scheduled, because the number of transition probabilities that need to be calculated increases exponentially. Hence, the model-user should be aware when multiple coupled operations in a project are scheduled or when a higher Markov chain order is applied.

Furthermore, the workability percentage per operation is validated with the RMSE, which resulted in very low values, concluding that the model preserves the workability percentage accurately. The persistency per operation is validated with the two-sample Kolmogorov-Smirnoff test, which was not accepted for all operations. For the not-accepted operations the model generated too optimistic weather windows compared to the hindcasted weather windows. But this overestimation can be controlled to some extent by increasing the Markov chain order, as illustrated in Section 2.5. The increase of the Markov chain order should be limited, else the model will replicate the hindcasted data. The slight overestimation in persistency is acceptable in terms of influence on downtime.

Table 3

Planned installation cycle with the net durations and operational limits of a hypothetical wind turbine foundation installation project [26].

No.	Description	Operational limit	Net duration
1	Load up to 5MPs and TPs from quay wall onto installation vessel	$U \leq 12 \text{ m/s}$	19 h
2	Sail to project site & jack-up	$H_s \leq 2,5 \text{ m}, U \leq 13 \frac{\text{m}}{\text{s}}$	19 h
3a	Installation (5x per cycle)		
	Handle MP and place in gripper frame	$H_s \leq 1 \text{ m}, U \leq 13 \frac{\text{m}}{\text{s}}$	4 h
3b	Pile MP & remove hammer		4 h
4a	Place TP	$H_s \leq 1,5 \text{ m}, U \leq 13 \frac{\text{m}}{\text{s}}$	4 h
4b	Grout TP	$H_s \leq 2 \text{ m}$	4 h
5	On-site relocation & jack-up	10 m keel clearance	1 h
6	Sailing back to port	$H_s \leq 2,5 \text{ m}, U \leq 13 \frac{\text{m}}{\text{s}}$	17 h

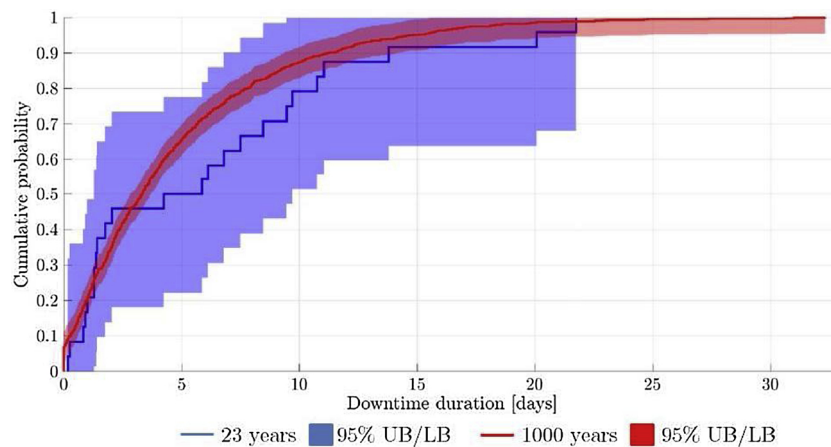


Fig. 10. The cumulative probability distributions of the downtime on the Tasman sea project with January-1 as start date. Downtime durations in blue are calculated with the original (23 years) and in red are calculated with the synthetic (1000 years) datasets [27].

4. Discussion

This research introduces a discrete-time 2-state-Markov model to stochastically simulate sequential operations in a marine project. The proposed model shows promising results for analyzing the downtime risk of marine projects. Especially for large cyclic projects more accurate downtime estimations are possible, since the variation in project duration of the current simulation methods is larger than for smaller projects. However, it should be kept in mind that the hindcasted data used as input for the downtime estimation is assumed to represent reality, which is not certain. If in the hindcasted data a once in 1000-year storm occurred, then the model will treat it as it happens every 20 years (if the dataset length is 20 years).

This model only incorporates downtime related to weather conditions, while in reality more uncertainties can influence the project downtime: the operational limit may be too restrictive, the net durations depend on the available crew and their learning curves, breakdowns of the equipment can occur. Additionally, the significant wave height is used in the calculations, but every single wave can be different. And in the end, the captain or project manager has the final responsibility on which decision is made. It is very complex to take all of these factors into account in a simulation model. Nevertheless, the proposed model makes a first step to at least quantify the caused delay by weather conditions. Therefore, it is expected that the proposed strategy can help tender teams to make decisions about the project duration they will put in tender.

In this paper only the simulation uncertainty is quantified, but in order to get a complete picture of the total uncertainty situation it is recommended to quantify the model and parametric uncertainties as well.

5. Conclusions

For offshore marine projects it is important to identify and map the uncertainties related to weather conditions. These weather conditions may delay certain operations, which goes hand in hand with additional financial costs. Current models to estimate marine project durations are making use of hindcasted data to perform their simulations: for each year in the data one realization of project duration is found. The upside

is that this metocean data is ‘real’ data (as if these project durations actually happened in the past), but there is a high simulation uncertainty in the prediction because it is made based on a small (generally ≤ 35 years) number of samples. A new model was proposed which is able to produce synthetic marine projects with linked Markov chains based on statistics of the metocean data near the project location.

In this paper the model was applied to a location in the Tasman sea with a hypothetical project (Table 3). It is demonstrated that the simulation uncertainty in the bandwidth is significantly reduced, as an unlimited number of years of project simulations can be generated. Also, the workability percentage is well preserved in the synthetic data compared to the original data. The persistency of sequential workable time steps is a slightly less well preserved, but this can be controlled to some extent by increasing the Markov chain order. Also, the dependencies between operations are well incorporated in the model and show promising results. It is recommended to consider the model at more locations with different projects, as it could lead to more general conclusions.

Any project and any location can be used to estimate the downtime in this model. Further application to other projects is recommended and the optimal settings of the model (e.g. choice of homogeneity and Markov chain order) will depend on local conditions. Optimal model settings can be determined with the homogeneity and order test, which influence the parametric and model uncertainty, and these should be minimized.

The added value of the model regarding downtime analysis in general, is that there was no stochastic model in current literature to simulate marine projects. The proposed model is an innovative and realistic way of modelling, because of the use of concepts as persistency, influence periods, n-order Markov chain, etcetera. Thus, it a promising alternative for existing project planning tools.

Acknowledgements

The contribution of Oswaldo Morales – Napoles and George Leontaris to this research is gratefully acknowledged. We thank Boskalis for supplying metocean data to perform the simulations.

Appendix A

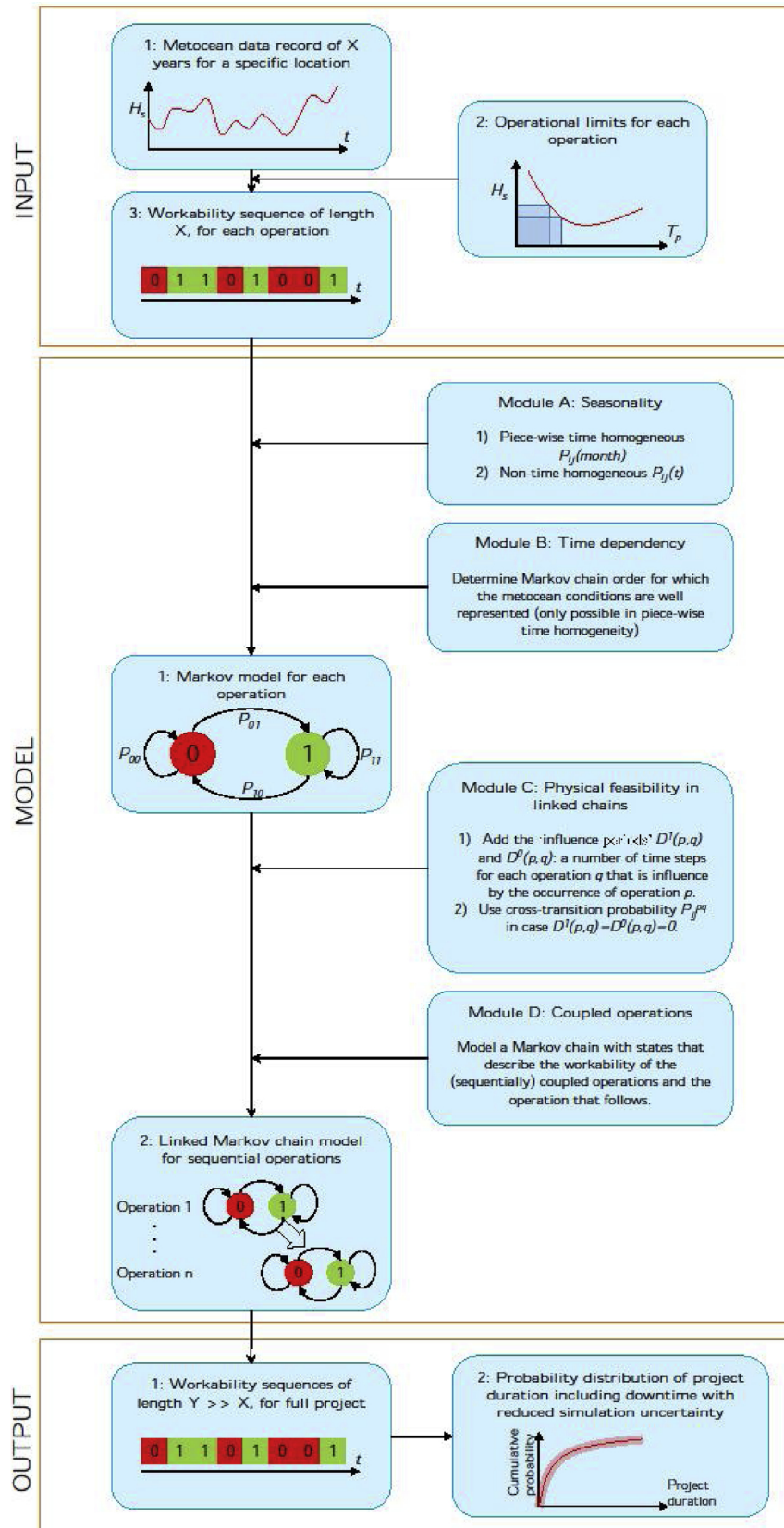


Fig. A1. Breakdown of proposed model for marine projects [26].

References

- [1] R. van der Wal, G. de Boer, Downtime analysis techniques for complex offshore and dredging, *Am. Soc. Mech. Eng.* (2004) 93–101.
- [2] C. Graham, The parameterization and prediction of wave height and wind-speed persistence statistics, *Coast. Eng.* 6 (1982) 303–329.
- [3] S. Kuwashima, N. Hogben, The estimation of wave height and wind-speed persistence statistics, *Coast. Eng.* 9 (1986) 563–590.
- [4] R.T. Walker, J. van Nieuwkoop-McCall, L. Johanning, R.J. Parkinson, Calculating weather windows: application to transit, installation and the implications on deployment success, *Ocean. Eng.* 68 (2013) 88–101.
- [5] M. O'Connor, T. Lewis, G. Dalton, Weather window analysis of Irish west coast wave data with relevance to operations & maintenance of marine renewables, *Renew. Energy* 52 (2013) 57–66.
- [6] K. Anastasiou, C. Tsekos, Operability analysis of marine projects based on Markov theory, *Appl. Ocean. Res.* 18 (1996) 329–352.
- [7] F. Arena, V. Laface, G. Malara, A. Romolo, Estimation of downtime and of missed energy associated with a wave energy converter by the equivalent power storm model, *Energies* 8 (10) (2015) 11575–11591.
- [8] D. Martins, G. Muraleedharan, C. Guedes Soares, Weather window analysis of a site off Portugal, *Maritime Technology and Engineering*, Taylor & Francis Group, London, 2015, pp. 1329–1338.
- [9] M. Mathiesen, Estimation of wave height duration statistics, *Coast. Eng.* 23 (1994) 167–181.
- [10] K. Anastasiou, C. Tsekos, Persistence statistics of marine environmental parameters from Markov theory, Part 1: analysis in discrete time, *Appl. Ocean. Res.* 18 (4) (1996) 187–199.
- [11] A. Ang, A. Chaker, J. Abdelnour, Analysis of activity networks under uncertainty, *J. Eng. Mech. Div.* 101 (1975) 373–387.
- [12] J. Feuchtwang, D. Infield, Offshore wind turbine maintenance access: a closed-form probabilistic, *Wind. Energy* 16 (2013) 1049–1066.
- [13] J. de Wilde, A. van Dijk, J. van den Berg, J. Dekker, Direct time domain downtime assessment for lng operations using computer cluster, *The Nineteenth International Offshore and Polar Engineering Conference* (2009).
- [14] Y. Kikuchi, T. Ishihara, Assessment of weather downtime for the construction of offshore wind farm by using wind and wave simulations, *Journal of Physics: Conference Series* 753 (2016).
- [15] M. Morandau, R.T. Walker, R. Argall, R.F. Nicholls-Lee, Optimisation of marine energy installation operations, *Int. J. Mar. Energy* 3 (2013) 14–26.
- [16] J.N. Hall, Use of risk analysis in North Sea projects, *Int. J. Proj. Manag.* 4 (1986) 217–222.
- [17] G. Leontaris, O. Morales-Nápoles, A.R. Wolfert, Probabilistic scheduling of offshore operations using copula based environmental time series—An application for cable installation management for offshore wind farms, *Ocean. Eng.* 125 (2016) 328–341.
- [18] S. Fouques, D. Myrhaug, F.G. Nielsen, Seasonal modeling of multivariate distributions of metocean parameters with application to marine operations, *J. Offshore Mech. Arct. Eng.* 126 (2004) 202–2012.
- [19] J.A. Bowers, G.I. Mould, Weather risk in offshore projects, *J. Oper. Res.* (1994) 409–418.
- [20] B. Hagen, I. Simonsen, M. Hofmann, M. Muskulus, A multivariate Markov weather model for O&M simulation of offshore wind parks, *Energy Procedia* 35 (2013) 137–147.
- [21] W. Xie, B.L. Nelson, R.R. Barton, Statistical Uncertainty Analysis for Stochastic Simulation, Doctoral dissertation Northwestern University, 2014.
- [22] D. Dee, S. Uppala, A. Simmons, P. Berrisford, P. Poli, S. Kobayashi, et al., The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.* 137 (2011) 553–597.
- [23] V. Monbet, P. Ailliot, M. Prevosto, Survey of stochastic models for wind and sea state time, *Probabilistic Eng. Mech.* 22 (2007) 113–126.
- [24] P. van Gelder, Statistical Methods for the Risk Based Design of Civil Structures, Delft University of Technology, Faculty of Civil Engineering and Geosciences, 2000 PhD thesis.
- [25] K.S. Hickmann, J.M. Hyman, S.Y. Del Valle, Quantifying Uncertainty in Stochastic Models, arXiv preprint arXiv:1503.01401, 2015.
- [26] J. Rip, Probabilistic Downtime Analysis for Complex Marine Projects, Master's Thesis TU Delft, 2015.
- [27] W. Bruijn, Probabilistic Downtime Analysis for Complex Marine Projects, Master's Thesis TU Delft, 2017.
- [28] W. Ching, K.N. Michael, Markov Chains. Models, Algorithms and Applications, (2006).
- [29] T.W. Anderson, L.A. Goodman, Statistical inference about Markov chains, *Ann. Math. Stat.* (1957) 89–110.
- [30] S. Ross, Introduction to Probability Models, Academic press, 2014.
- [31] B. Rajagopalan, U. Lall, D.G. Tarboton, Nonhomogeneous Markov model for daily precipitation, *J. Hydrol. Eng.* 1 (1996) 33–40.
- [32] B. Tan, K. Yilmaz, Markov chain test for time dependence and homogeneity: an analytical and empirical evaluation, *Eur. J. Oper. Res.* 137 (3) (2002) 524–543.
- [33] R. Castro, Lecture Notes in Applied Statistics (tu Eindhoven), Lecture 1 - Introduction and the Empirical, (2013).
- [34] B.P. Flannery, W.H. Press, S.A. Teukolsky, W. Vetterling, Numerical Recipes in C, Press Syndicate of the University of Cambridge, New York, 1992.