Omega 66 (2017) 58–78

Contents lists available at ScienceDirect

Omega

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

# Optimised scheduling for weather sensitive offshore construction projects ☆

L.-P. Kerkhove [a], M. Vanhoucke [a,b,c,*]

[a] Faculty of Economics and Business Administration, Ghent University, Ghent, Belgium
[b] Technology and Operations Management Area, Vlerick Business School, Ghent, Belgium
[c] UCL School of Management, University College London, London, UK

ARTICLE INFO

*Article history:*
Received 13 April 2015
Accepted 27 January 2016
Available online 6 February 2016

*Keywords:*
Offshore construction
Weather simulation
Stochastic scheduling
Project scheduling

ABSTRACT

The significant lead times and costs associated with materials and equipment in combination with intrinsic and weather related variability render the planning of offshore construction projects highly complex. Moreover, the way in which scarce resources are managed has a profound impact on both the cost and the completion date of a project. Hence, schedule quality is of paramount importance to the profitability of the project. A prerequisite to the creation of good schedules is the accuracy of the procedure used to estimate the project outcome when a given schedule is used. Because of the systematic influence of weather conditions, traditional Monte Carlo simulations fail to produce a reliable estimate of the project outcomes. Hence, the first objective of this research is to improve the accuracy of the project simulation by creating a procedure which includes both uncertainty related to the activities and an integrated model of the weather conditions. The weather component has been designed to create realistically correlated wind- and weather conditions for operationally relevant time intervals. The second objective of this research is to optimise the project planning itself by using both general meta-heuristic optimisation approaches and dedicated heuristics which have been specifically designed for the problem at hand. The performance of these heuristics is judged by the expected net present value of the project. The approach presented in this paper is tested on real data from the construction of an offshore wind farm off the Belgian coast and weather data gathered by the Flanders Marine Institute using measuring poles in the North Sea.



## 1. Introduction

This paper presents a comprehensive methodology for the proactive scheduling of capital intensive projects which are sensitive to weather conditions. This is done by combining heuristic optimisation techniques and weather simulation models which are fitted to data gathered on the North Sea off the Belgian coast. The construction of an offshore wind farm will be used as an archetypical example of a capital-intensive and weather-sensitive project.

The significant lead times and costs associated with materials and equipment in combination with the significant stochasticity introduced by varying weather conditions make the planning of offshore construction projects highly complex. Moreover, substantial economical gains are made when the planning of such projects can be optimised. The goal of this paper is to improve upon the current best practices for creating schedules for such projects.

This paper presents a novel weather simulation model which is capable of generating correctly correlated wind- and wave information. Moreover the paper also presents a novel model for the project and the associated schedule. Both these models include a greater degree of realism than the current state-of-the-art, resulting in more accurate estimates of the value of specific scheduling strategies.

Finally, this model also introduces and tests several scheduling heuristics which aim to maximise the expected net present value derived from a project. These heuristics are tested on both a specific case study in offshore wind and generated projects of varying sizes.

## 2. Literature

This study combines techniques from project scheduling literature with insights from weather modelling, specifically stochastic weather generation models. This was motivated by the

☆This manuscript was processed by Associate Editor Kis

* Corresponding author at: Faculty of Economics and Business Administration, Ghent University, Ghent, Belgium.

http://dx.doi.org/10.1016/j.omega.2016.01.011
0305-0483/

arguments made by Regnier [35], who stated that recent advances in meteorological sciences have created substantial opportunities for quantitative modelling. Specifically, the OR/MS community could make significant contributions by tailoring meteorological products to the users' decision contexts. Incidentally, this is the main objective of this research.

Section 2.1 provides an overview of the literature on weather simulation models. Given the objectives of this research, the focus lies on models focussing on generating realistic weather patterns, rather than models with more meteorological objectives. This literature is the basis for the weather model presented in Section 4.2.1.

The second part of the literature review focusses on the contributions from the field of project management (Section 2.2). This encompasses recent advances in dealing with uncertain activity durations, resource constraints and non-regular objectives which more accurately represent the financial implications of project scheduling. This foundation is then used to construct a model for large-scale offshore construction projects in Section 3.1. As was noted earlier, these theoretical foundations are supplemented with information from offshore wind projects (see Section 5).

### 2.1. Modelling weather influences

Accurately estimating the weather has been a goal in operations management for over a century. The majority of models which have been created for this purpose have focussed on predicting precipitation, which is of major importance to agriculture and many other industries Wilks and Wilby, [51].

For offshore construction there are two key weather factors which determine whether working is possible or not: the significant wave height and the average wind speed Graham, [16]. Depending on the nature of an activity, an acceptable threshold for both these dimensions can be defined. If this threshold is exceeded, the activity has to be interrupted and resumed when weather conditions have improved. Such interruptions can of course be highly detrimental to project performance.

#### 2.1.1. Types of weather models

Two different kinds of weather models can be distinguished: weather simulation models and weather forecasting models. Simulation models can be viewed as a highly complex random number generators whose outputs resemble weather conditions at a certain location. However, these random sequences are in no way meant to be a forecast of what is going to happen; i.e. the weather pattern remains completely fictitious, but is assumed to have similar properties to weather at the location. Weather forecasting models on the other hand operate from a deterministic point of view and try to make short term predictions of what the weather will be like Wilks and Wilby [51].

Both simulation and forecasting models are highly useful when managing offshore construction models. A good forecasting model can help project management with day-to-day operational decisions, allowing them to delay work when weather conditions are expected to be suboptimal. Weather simulation on the other hand can help to optimise the scheduling of a project in order to make it more robust with regards to weather influences. This robustness is highly important since offshore construction projects are highly capital intensive due to the use of expensive machinery. Impromptu leasing of this machinery is not possible, hence it is highly important to make accurate predictions of when this machinery will be needed as well as accurate estimates of the duration for which this machinery will be needed.

Both simulation and forecasting models can be used for either meteorological or economical purposes. Where the focus of the former lies on providing an explanation for a physical phenomenon, the latter is interested in the impact of weather on a certain process or activity. This distinction obviously has a major impact on the way the model is constructed. Models wishing to explain meteorological phenomena usually use much smaller time steps (less than one hour) than models used for economical purposes, which almost always use time steps between one and six hours, and in some cases even more Monbet et al. [29]. This different time step sizes also impact the choice of modelling techniques. An example of this are Markov Chains which have been shown to be less adequate for weather models taking time steps of less than 40 min [8].

Monbet et al. [29] also highlight the importance of the trade-off between a model's ease of use and the accuracy of a model. A model's ease of use can be expressed using a number of qualitative criteria. Firstly, the robustness to the data source which may contain a considerable number of missing values. Secondly, the ease of use also depends on the robustness to the process itself. Some models are easier to use for a broader set of processes than others, especially for weather models where complex correlation between different elements may occur. Finally, the mathematical complexity and the time needed to implement such a model are also important. A model should not be needlessly complex thereby squandering scarce resources for its implementation when a simpler model would have sufficed. The accuracy of a model signifies the degree to which the statistical properties of the model correspond to the statistical properties of the original data. This includes various statistical properties such as the average, standard deviation, seasonality and autocorrelation.

#### 2.1.2. Dealing with non-stationary data

Many traditional forecasting and simulation models such as Markov chains and autoregressive models are valid under the assumption that the underlying process is stationary.[1] Most weather parameters however are subject to strong seasonal influences. Moreover, some weather variables such as wind can differ depending on the time of day. To deal with this issue many authors have proposed methods to transform this non-stationary data into stationary or quasi-stationary processes.

There are two ways in which this issue can be resolved [29]. The first method uses statistical techniques to remove the seasonal component from the data, the resulting series is then stationary and the aforementioned techniques can then be used [42,43,29]. A potential issue here is that these statistical techniques may not succeed in removing all the non-stationarity from the data. An alternative method is fitting a separate model for each season, assuming that stationarity can be assumed within a certain season. The most important drawback of this method is the need for substantially more data since a model has to be estimated for each season. Furthermore, this method also introduces artificial breaks in the data, which are of course not present in reality.

#### 2.1.3. Modelling techniques

This section presents a short overview of weather modelling techniques, discussing their key advantages and drawbacks. These approaches are grouped using the categorisation used by Monbet et al. [29], who distinguish Gaussian-based parametric techniques, non-Gaussian based parametric techniques and non-parametric techniques.

*Gaussian-based parametric techniques*

Gaussian-based parametric techniques remove seasonality from data using statistical techniques, such as Box–Cox and Fourier transformations [10]. Once this transformation is

[1] A stationarity process is a process whose joint probability does not change when the process is shifted in time.

complete, a wide variety of techniques such as autoregressive (AR) models or translated Gaussian processes are fitted to the deseasonalized pattern. Because all weather data is used to estimate a single model, relatively little data is needed to estimate parameters when compared to techniques such as Markov chains which typically split the available data into seasons which are modelled separately. Moreover, Gaussian-based techniques are capable of identifying trends in potentially noisy weather data. Nevertheless, these techniques fall short when multiple weather components (i.e. wind, wave height, precipitation, etc.) have to be modeled simultaneously [29].

*Non-Gaussian-based parametric techniques*

Several other parametric techniques have been used to model weather. Finite state space Markov chains and their variants are undoubtedly the most widely adopted for offshore weather simulation purposes. Rothkopf et al. [36] present a first-order Markovian model which defines a number of sea states and define a transition matrix based on the locality and season in which operations are carried out. The authors also note that a more accurate representation of the weather conditions can be obtained by increasing the number of states considered in the Markov weather model, insofar as the quantity of available observations allows this.

The persistence of good and bad weather spells is of paramount importance to offshore construction projects. Because of this, a number of authors have focussed on modelling this persistence and incorporating it in simulation models. Graham [16] describes how weather simulations can be made to take into account persistence by using a transition matrix in combination with two-parameter Weibull distributions which model persistence of a certain state.

Bowers and Mould [7] specifically investigate weather risk for offshore construction projects. The model used to simulate weather conditions combines a Markov chain analogous to the one presented in Graham [16] with a gamma distribution. In this model six different transition matrices are used to represent the six main seasons in which the simulation operates. Each of these transition matrices defines the possibility of transferring between one of 5 possible states. Because the persistence behaviour of the Markovian model in the first (calmest) of the five states did not correspond to the historic data, a gamma distribution was used to randomly draw the persistence upon arrival in the first state. Another example of the use of Markov chains to simulate wave height is presented by Norlund et al. [31], who used the technique to create robust and more environmentally friendly schedules for offshore supply vessels.

Markov chains have also been used to simulate wind rather than significant wave heights. A study by Sahin and Sen [37] shows that even a very simple first order Markov chain with eight states can attain an adequate level of accuracy. Ettoumi et al. [12] used Markov chains in combination with a Weibull distribution to model both wind speed and wind direction. For the wind speed each of the discrete states of the transfer matrix is linked to a certain Weibull distribution, from which the exact wind speed is then drawn at random. Second order Markov chains were used by Shamshad et al. [40] to make short-term predictions of wind speed. The authors noted that there is a slight improvement in terms of autocorrelation when using a second order rather than first order Markov chain model.

Brokish and Kirtley [8] make an important observation with regard to the use of Markov Chains used to simulate wind power, namely that this technique is often unable to match the autocorrelation of the original distribution when used for time steps which are shorter than 40 min. Hence, this technique should only be used for situations with longer time steps.

Hidden Markov models have also been used successfully to model weather conditions, one example of this is the Markov Switching Autoregressive model (MS-AR) which was proposed by Ailliot et al. [2].

Another parametric technique which is often used for weather prediction weather is an artificial neural network (ANN). More and Deo [30] used artificial neural networks to forecast wind speeds in coastal regions and found that these artificial intelligence techniques outperform traditional ARIMA techniques. The usefulness of this technique for weather prediction has also been illustrated by Arena and Puca [3] who have used this technique to reconstruct significant wave heights based on correlated data from other measurement stations. Makarynskyy et al. [24] also used ANN to predict wave height and period in order to deliver a warning signal up to 24 h in advance. More recently Mandal and Prabaharan [25] have also used the ANN technique to make short term predictions for significant wave height.

Nonlinear autoregressive models are another technique which may be used rather than the traditional Gaussian techniques. Due to the fact that the underlying process in inherently nonlinear, nonlinear models can improve the estimation of the higher statistical moments such as skew and kurtosis [39].

*Non-parametric techniques*

The use of non-parametric techniques is considerably less widespread than that of parametric techniques. One technique which is sporadically used for making bootstrap estimations is resampling [29]. This technique consists of a nearest neighbour heuristic with a random component to create weather patterns which are similar to the patterns observed in the original data. The advantage of this technique being that no assumptions have to be made regarding the stationarity of the underlying data. Hence, it is unneeded to construct separate models for different seasons, nor is it required to subject the data to transformations to remove seasonal patterns. The key disadvantage of this technique is the significant risk of overfitting, resulting in weather patterns which resemble the observed data too closely.

### 2.2. Project scheduling for offshore construction projects

Offshore construction projects present one of the most challenging operational environments. Within this environment, the operational strategy can significantly impact the profitability, robustness and the environmental impact of a project [31], and even prevent the loss of lives [33,34]. This section highlights the specific challenges faced when scheduling offshore construction projects. This is done by presenting a concise review of relevant project scheduling literature. This literature forms the basis for the project model presented in Section 3.1.

First and foremost, offshore projects are highly susceptible to adverse weather conditions. Rothkopf et al. [36] and Graham [16] highlight the importance of extended periods of good weather for the execution of certain activities, as well as the significant startup cost which is incurred when an activity has to be interrupted due to harsh weather conditions. Hence, the project manager will try to optimise his project in light of the weather risks (s)he is facing. Bowers and Mould [7] have analysed how this can be done by either delaying the start of the project, or by using more expensive equipment which is less susceptible to extreme weather conditions. Obviously, many other strategies can also be used to improve the expected performance of a project.

Traditional project scheduling literature usually analyses risk in a more general way, by considering activity costs and/or durations to be stochastic rather than deterministic. Although the majority of project scheduling research still uses the assumption that activity durations are deterministic [19], several authors have proposed methods for scheduling whilst taking stochasticity into account. An overview of these methods is presented in Section 2.2.1.

A second challenge faced in the scheduling of offshore construction projects is the planning of expensive resources needed to perform the activities. The chartering of jack-up barges, transport ships, dumping vessels, crane vessels, etc. [13] is the key cost in the majority of offshore construction projects. Especially since many of these vessels have to be reserved up to two years beforehand, and significant mobilisation and demobilisation costs are incurred [38]. These issues can be related to resource constrained scheduling, as well as resource renting problems. These concepts are discussed in further detail in Section 2.2.2.

The combination of weather sensitive activity durations and substantial variable costs associated with activities results in a significant financial risk. Hence, it makes sense to use a financial objective, rather than makespan-based targets. More specifically, the main interest of the project manager in this case is the net present value of the complete project. The default formulation of this problem assumes that a fixed positive or negative cashflow is associated with the start of each activity, the goal being to speed up cash inflows and delay cash outflows. This basic formulation has been extended by considering time-dependent rather than static cashflows. Section 2.2.3 presents an overview of the most relevant publications which analyse this issue.

#### 2.2.1. Dealing with uncertainty

The techniques used for dealing with uncertainty that typifies projects in progress are frequently summarised using the term *dynamic scheduling* [49]. Within dynamic scheduling, three basic approaches can be distinguished: a reactive strategy, a proactive strategy or a mixed strategy [14]. The literature review on stochastic scheduling by Herroelen and Leus [19] provides even more detail and identifies six techniques which can be used to deal with stochasticity in a project scheduling environment.

Given that resources needed for the execution of offshore construction activities have to be reserved long beforehand, the focus of this research lies on the proactive approach. More specifically the construction of robust schedules, which are capable of dealing with the impact of adverse weather conditions in a satisfactory manner.

Practically, the creation of schedules for stochastic project scheduling problems is most frequently done by defining priority lists. Several different methods for creating these lists exist, each has advantages and disadvantages. Research by Stork [44] has shown that linear restrictive policies can be expected to perform best in situations where resource restrictions (as defined by minimal forbidden sets) are limited. When this is not the case, and complex resource restrictions are present in the project, it is better to use activity-based priority rules. A key downside of this simple approach is that it is not possible to insert buffers at strategic moments in time when it may be ill-considered to proceed immediately with the next activity on the priority list.

Another approach which resolves this issue has been analysed by Ke and Liu [22], who use a genetic algorithm in combination with a simulation procedure in order to determine the earliest start times of activities. This algorithm is tested on the three basic approaches for stochastic optimisation: the maximisation of the objective, the maximisation of the objective value attained with a certain probability $\alpha$, or the maximisation of the probability of obtaining a certain objective function value. The latter two scenarios being of great significance when the aim of the project manager is protecting the downside risk associated with the project.

The method of improving schedule robustness by inserting time buffers at critical points in the schedule is studied more explicitly by Herroelen and Leus [18]. The authors in this study analyse how the robustness of a schedule can be maximised for a given project deadline. The concept *float* is used to express the magnitude of the buffer inserted between activities. Three different types of float are identified for an activity: the *pairwise float* is equal to the amount of float between two activities which are connected to each other by a precedence relation. This float can be calculated by subtracting the planned finish time of the predecessor from the planned starting of the successor. The two other float metrics are derivations from this basic concept. The *free float* is the maximal delay of an activity, without affecting the start of it successors. Note that this metric takes all predecessors and successors into account. The *total float* is similar, but now the successors of the activity are assumed to start at their latest possible times. The origins of this technique can be traced back to earlier studies in machine and job-shop scheduling [26], where it was used to maximise a deterministic metric representing the robustness of a schedule.

Such deterministic proxies for robustness are hard to use in the situation studied in this paper, where the stochasticity depends on a complex external process. An alternative for these deterministic metrics is presented by Huang and Ding [21], who use a genetic algorithm combined with stochastic simulations to estimate the outcome of a specific strategy.

#### 2.2.2. Dealing with resources

The resource constrained project scheduling problem, commonly known as the *RCPSP*, is the most basic way of dealing with resources in a project environment Hartmann and Briskorn, [17]. A variation on this approach is the resource renting problem, which does not only include renewable resources, but also takes into account the cost of renting these resources. Specifically, the costs associated with mobilisation and demobilisation of resources can inflate the total cost of a project to a project owner. This problem class was originally introduced by Nübel [32], and more complex heuristic solution techniques have since been presented by Ballestín [4,5]. Extensions including more complex cost structures and the possibility of scheduling in overtime have also been presented in recent years [47,23]. These more complex cost structures have a number of analogies with the chartering process used in offshore construction projects.

However, to the best of our knowledge literature on this problem has only studied the resource renting problem under the assumption of deterministic activity durations. Nevertheless, other authors have investigated methods of dealing with resources in a stochastic project setting. The CC/BM methodology presented by Goldratt [15] is the most widely known methodology to deal with similar situations.

A more specialised method for doing this is presented by Trietsch [46], who uses *gates* to represent the moment in time when a resource has been made available for a certain activity. This methodology also defines two costs associated with these gates: a *holding cost*, which is incurred when an activity has to wait for resources, and a *shortage cost*, which is incurred when resources have already been made available, but the activity is still waiting for one or more predecessors to finish. A recent paper by Bendavid and Golany [6] also uses this concept of gates in combination with a solution technique based on the cross entropy methodology in order to schedule a project with stochastic (exponential) activity durations.

#### 2.2.3. Maximising financial returns

When cashflows can be determined for each individual activity, it is possible to use the project's net present value (*NPV*) as an objective measure of the outcome of the project, and by extension the comparative quality of different scheduling strategies. Traditionally, this is done within a deterministic setting where both the activity durations and the magnitude of the cashflows are assumed to be set with 100% accuracy beforehand. The problem described in this paper however has a very pronounced stochastic component, which influences both the timing and the magnitude of the cashflows associated with the different activities. These variations are caused by stochasticity due to weather conditions, as well as the intrinsic variability of the activities. Although this specific scenario has not yet been investigated in existing literature, it can be linked to earlier problem formulations which have considered the implications of time-dependent cashflows. The remainder of this section focusses on authors which have investigated net present value optimisation while assuming some degree of variability of the cashflows associated with activities.

One such study is presented by Etgar et al. [11], who investigate how projects consisting of activities with cashflows that increase or decrease over time can be scheduled to maximise net present value. An exact solution procedure for this problem was later presented by Vanhoucke et al. [50]. A variation on this problem where certain regions are defined wherein the cashflows remain constant is presented by Achuthan and Hardjawidjaja [1].

The preceding authors all assumed that cashflows linked to activities could either increase, decrease or remain constant over time. However, for the problem studied in this paper it is possible that cashflows linked to activities display seasonal patterns, as variable costs are heavily influenced by weather conditions. This assumption was relaxed by Möhring et al. [28], who defined a specific cost based on the starting time of an activity, but the evolution of this cost had the potential of fluctuating freely over time. In a later paper Möhring and Schulz [27] discovered that this problem formulation can be solved using fast minimum-cut operations.

The fluctuations in activity cashflows for the problem studied in this paper are caused by stochastic variations in activity duration, depending on the time when the activity is executed. This is not completely identical to the problems studied in the preceding papers, which assumed that the cashflow itself is directly affected by the moment in time. Nevertheless, formulations which use time-dependent processing times have already been investigated in the context of machine scheduling [9].

One key limitation of all preceding studies is that they consider time-dependence within a deterministic rather than stochastic setting, something which some of these authors had already identified as a key limitation and possible avenue for future research [11]. Sobelet al. [41] depart from this deterministic setting and maximise *NPV* for a project with stochastic activity durations using a reactive priority list approach. Such reactive strategies are, however, not adequate for the problem studied in this paper. The reason for this being the significant lead times associated with both consumable and renewable resources required to execute activities. Hence, the focus of this research is to design a proactive approach in order to create robust solutions.

A recent paper by Huang and Ding [21] presents such a proactive approach in a stochastic problem environment. Specifically, this is done through the combination of stochastic optimisation and a genetic optimisation algorithm. Nevertheless, the problem considered by Huang and Ding [21] uses a cost objective rather than the net present value. Moreover, solely the duration of activities is assumed to be subjected by stochasticity – the actual cashflows are assumed to remain unaltered.

### 2.3. Concluding remarks

This review of literature has shown that the proactive scheduling of offshore construction projects can be improved considerably by using recent advances in meteorological models to improve model accuracy. This premise is motivated by the research by Regnier [35], who stated that tailoring these meteorological techniques to specific OR problems is an area where significant contributions have yet to be made.

Based on the review of existing weather simulation models in Section 2.1, a Markovian model capable of simulating both wave and wind conditions is proposed in this paper. This capability to create realistic simulations of both weather parameters which impact the construction activities is he main motivation for choosing this type of model. The key disadvantage of such a model being that it needs a considerable amount of data to be fitted adequately. Nevertheless, since large datasets containing both wind and wave information are available, this was not an issue for this research. Similarly, the limitations regarding the minimal step size imposed by this type of model were not operationally relevant for the offshore construction projects investigated in this research.

The overview of the state of the art in related project management literature (Section 2.2) revealed that although attention has increased in recent years, a lot of potential remains in the area of stochastic project scheduling. Specifically, the maximisation of net present value in projects where the cashflows are influenced by systematic stochasticity has never been investigated. Building on works of earlier authors who have identified the major challenges in offshore construction projects [36,16], this research proposes and tests the performance of novel techniques.

## 3. Problem definition

This section introduces the components of the problem investigated in this research. First, a mathematical representation of the scheduling problem is presented in Section 3.1. Next, the stochastic influences on this model are discussed in Section 3.2. Finally, the problem is illustrated by means of a theoretical example in Section 3.3.

### 3.1. The scheduling problem

The goal of this problem is to find the optimal 'gate times' (i.e. the time at which resources are made available for a specific activity) in order to maximise the net present value (*NPV*) of the complete project. These gate times are not only a theoretical representation of the schedule, but are actually required in order to guarantee the availability of the offshore construction vessels before the start of the project. The reasons for adopting the net present value as the objective are the following:

1. Offshore construction projects are highly capital intensive because of the high material cost and the need to charter large construction vessels. Hence, the objective used should take into account the magnitude and timing of these financial flows.
2. Due to the conditions in which these projects are executed, the timing and magnitude of cashflows can be volatile. Using the net present value allows for a more objective comparison of such volatile scenarios when compared to an optimisation of the total cost.
3. Linked to the preceding point is the fact that because of the inherent volatility, investors will require a larger return on investment for the project. Hence, the discount rate is larger than it would be for other projects. Again, this would mean a large discrepancy between the total costs incurred and the net present value of the project.
4. Because of the significant length (often multiple years) of offshore construction projects, it is required to discount both the cash outflows and the cash inflows to get an accurate image of the profitability of the project.

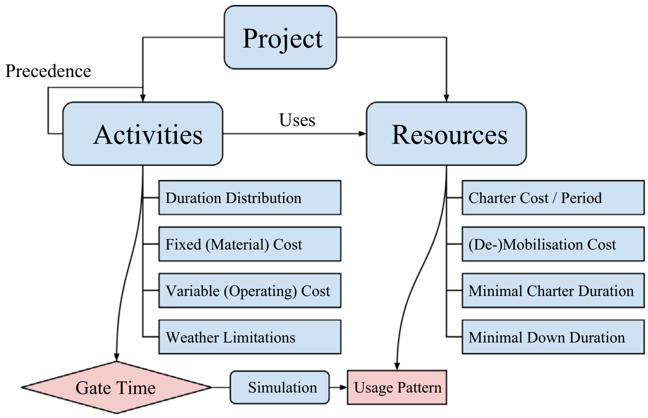


**Fig. 1.** Schematic representation of the project model.

For projects with the aforementioned properties, there is a decided preference for the use of net present value as the opimisation objective [20]. Other objectives such as the minimisation of the project lead times are likely to result in reduced project profits.

Fig. 1 visualises the various elements of the project which are relevant to the optimisation. The project can be seen as a collection of interrelated activities and resources. The activities are interconnected by precedence relationships and are also linked to the resources they require to be executed. Within the context of offshore construction these resources are the construction vessels.

An activity is defined by its duration ($d_i$), its fixed and variable cost ($C_i^F$ and $C_i^V$) and the weather limitations under which it can operate. The fixed cost associated with an activity represents the cost of the materials which are used during the activity. For offshore wind projects, this constitutes (but is not limited to) the turbine itself, as well as its foundation and scour protection – naturally these are allocated to the respective activities. The variable cost is the operating cost incurred during the execution of the activity, this includes the fuel cost of operating the vessels as well as the costs for personnel operating at sea. However, this does not include the chartering cost of the vessel itself which has to be paid even when the vessel is temporarily not used but not decommissioned, this cost is linked to the specific resources. The activity's weather limitations are expressed in two dimensions: the maximal wave state and the maximal wind state at which the activity can be executed. Exceeding these thresholds will cause the activity to be temporarily halted until weather conditions improve. Note that it is of paramount importance to distinguish between these two dimensions for offshore operations since different types of activities will have a different degree of wind- and wave-sensitivity. For the ease of notation this is simplified in the deterministic model presented below, using the binary $\omega_{it}$ parameter to indicate if the weather threshold for activity $i$ is not exceeded in time period $t$.

Separate costs are associated with the resources used by the activities (i.e. the installation vessels): the charter cost per period ($\rho_r^V$), the mobilisation and demobilisation cost ($\rho_r^C$ and $\rho_r^D$), as well as the minimal duration of both the chartering ($\delta_{up}^{\min}$) and the downtime ($\delta_{down}^{\min}$). The basic cost of a vessel is the daily amount paid as rent, however at the start of the renting period the operator also has to pay to deploy the vessel at the desired location ($\rho_r^C$). Similarly, costs are also associated with the decommissioning of these vessels ($\rho_r^D$). These costs vary depending on the vessel type but are often equal to 4–8 times the associated day-rate of the vessel [38]. Related to these mobilisation and demobilisation charges are the minimal durations for both chartering and downtime, which are required by the company which leases the vessels to ensure better work continuity [48]. The minimal charter duration guarantees that the vessel will be chartered for a sufficiently long period. Similarly it is undesirable to decommission and recommission vessels with only very short downtimes, simply because it is unlikely that the vessel's owner can find a client for this small timeframe.

Because of this (de)mobilisation decision, the chartered resources are not simply viewed as a variable cost that is directly linked to an activity. Once a schedule is constructed, the demand for a certain vessel can be determined and the periods when the vessel needs to be chartered can be fixed. An important note here is that regardless of the fact that the acquisition of these resources has a very long lead time, the contractual agreements generally allow for a scheduling contingency, meaning that resources will not be deallocated until the work for which they have been chartered has been completed.

The schedule itself is defined by setting *gate times* for all activities, these times indicate when the resources to execute an activity will be made available. Hence, the time at which an activity can start is the maximum of this gate time and the latest finish time among its predecessors. This principle of gate times is similar to the methodology used by Trietsch [46] and Bendavid and Golany [6] (see Section 2.2.2). Practically, this means that the gate time is directly linked to the fixed cost associated with the activity. Hence, this fixed cost will always be incurred at precisely the gate time of the activity. The variable activity costs, as well as the costs associated with chartering vessels will be incurred when the activity is actually executed.

The simplest representation of the problem studied in this paper is a deterministic scheduling problem as described by Eq. (1) until Eq. (10). The mathematical model below assumes that both the duration of the activities and the weather pattern is completely deterministic, an assumption which will be relaxed in Section 3.2. The notation used in this scheduling optimisation model is summarised in Table 1.

$$\max NPV = \sum_{i=1}^{NI}\left[\frac{-C_i^F}{(1+ROI)^{g_i}} - \frac{C_i^V\cdot(f_i - s_i)}{(1+ROI)^{f_i}}\right] + \sum_{r=1}^{NR}\sum_{t=1}^{NT}\left[\frac{-\rho_r^V\cdot u_t^r}{(1+ROI)^t} - \frac{\rho_r^C\cdot v_t^r}{(1+ROI)^t} - \frac{\rho_r^D\cdot w_t^r}{(1+ROI)^t}\right] + \frac{\pi}{(1+ROI)^{f_{NI}}} \tag{1}$$

subject to:

$$s_i = \max\left[g_i, \max_{j\in P_i}(f_j)\right] \quad \forall i\in I \tag{2}$$

$$\sum_{t=s_i}^{f_i-1}\omega_{it} = d_i \quad \forall i\in I \tag{3}$$

$$u_{rt} = \sum_{i\,|\,s_i\le t<f_i}\omega_{it} \quad \forall r\in R \quad \forall t\in T \tag{4}$$

$$\varrho_{rt} \ge u_{rt} \quad \forall r\in R \quad \forall t\in T \tag{5}$$

$$\varrho_{r1} = v_{r1} \quad \forall r\in R \tag{6}$$

$$\varrho_{rt} = \varrho_{r(t-1)} + v_{rt} - w_{rt} \quad \forall r\in R \quad \forall t = 2,\dots,NT \tag{7}$$

$$w_{rt} \le M\cdot\left[1-\min\left(1, \sum_{t'=\max(1,t-\delta_{up}^{\min})}^{t} v_{rt'}\right)\right] \quad \forall r\in R \quad \forall t\in T \tag{8}$$

$$v_{rt} \le M\cdot\left[1-\min\left(1, \sum_{t'=\max(1,t-\delta_{down}^{\min})}^{t} w_{rt'}\right)\right] \quad \forall r\in R \quad \forall t\in T \tag{9}$$

**Table 1**
Overview of notation.

| Indices | |
|---|---|
| $i$ | Activity index: $i \in I = \{1, \ldots, NI\}$ |
| $r$ | Resource index: $r \in R = \{1, \ldots, NR\}$ |
| $t$ | Time interval index: $t \in T = \{1, \ldots, NT\}$ |
| *Decision variables* | |
| $g_i$ | Gate time of activity $i$ |
| *Assisting variables* | |
| $s_i$ | Starting time of activity $i$ |
| $f_i$ | Finish time of activity $i$ |
| $u_{rt}$ | Resource usage of $r$ at time $t$ |
| $v_{rt}$ | Commissioned amount of $r$ at time $t$ |
| $w_{rt}$ | Decommissioned amount of $r$ at time $t$ |
| $\varrho_{rt}$ | Amount of resources $r$ rented in period $t$ |
| *Stochastic parameters* | |
| $\omega_{it}$ | Binary $=1$ if suitable weather for act $i$ at $t$ |
| $d_i$ | Duration of activity $i$ |
| *Fixed parameters* | |
| $ROI$ | Discount rate per period |
| $C_i^F$ | Fixed cost of activity $i$ |
| $C_i^V$ | Variable cost of activity $i$ |
| $\rho_r{}^V$ | Variable chartering cost for resource $r$ |
| $\rho_r{}^C$ | Commissioning cost for resource $r$ |
| $\rho_r{}^D$ | Decommissioning cost for resource $r$ |
| $\pi$ | The cashflow associated with the completed project |
| $\mu_{ir}$ | Binary $=1$ if act $i$ requires resource $r$ |
| $P_i$ | Set of predecessors of activity $i$ |
| $\delta_{up}^{\min}$ | Minimal renting period required for resources. |
| $\delta_{down}^{\min}$ | Minimal downtime period required for resources |
| $M$ | A large number |

$$g_i, s_i, f_i \in \mathbb{N}_0 \varrho_{rt}, \quad u_{rt}, v_{rt}, w_{rt} \in \mathbb{N} \tag{10}$$

The main decision variable in the optimisation model is the gate time $g_i$, which represents the time period at which the resources for an activity have been made available. The NPV-objective is calculated by Eq. (1), which is a function of the various cashflows associated with the project. The fixed cost ($C_i^F$) represents the building materials required to perform an activity. Since the gate time $g_i$ represents the moment in time when these resources have been made available this is the discount factor for the fixed cost. As explained earlier, the variable costs ($C_i^V$) are incurred in all time periods as soon as the activity has actually started. These costs are assumed to be payable upon completion of the activity. The second part of Eq. (1) represents the costs associated with the resources (i.e. the chartering of the vessels). This includes the variable operating costs $\rho_r^V$ incurred when the resource is actually used, as well as the commissioning ($\rho_r^C$) and decommissioning ($\rho_r^D$) costs associated with the vessel. Finally, the positive cashflow associated with completing the project is also added to the NPV.

The earliest possible starting time of an activity is either the gate time $g_i$, or the time period after the finish of its latest predecessor, whichever is greater (Eq. (2)). The activity finishes ($f_i$) as soon as the number of time periods with acceptable weather conditions since the start of the activity is equal to the duration of the activity $d_i$ (Eq. (3)). In order to calculate the variable operating cost of the resources the variable $u_{rt}$ is introduced. This variable is equal to the resource use for resource $r$ during period $t$, and is assigned by considering all active activities during time period $t$ as is shown in Eq. (4). In order to calculate the commissioning and decommissioning costs correctly, the variable $\varrho_{rt}$ is introduced. This variable is equal to the available amounts of resource type $r$ during period $t$. Naturally, the available amount of resources must always exceed the amount of resources used (Eq. (5)). The evolution of the $\varrho_{rt}$ variable is tracked by the $u_{rt}$ and $w_{rt}$ which track the commissioned and decommissioned resources during each time period respectively. The correct values for these variables is enforced by Eqs. (6) and (7). Finally, Eqs. (8) and (9) enforce the minimal periods enforced on the commissioned and decommissioned periods as explained above.

In the next section stochasticity is introduced into the scheduling problem, prior to presenting a small theoretical example in Section 3.3 in order to familiarise the reader with the concepts which have just been introduced.

**Table 2**
Example project information.

| Activity property | 1 | 2 |
|---|---|---|
| Duration distribution | $U(9, 11)$ | $U(8, 12)$ |
| Fixed cost ($C_i^F$) | € 100 | € 175 |
| Variable cost ($C_i^V$) | € 5 | € 10 |

### 3.2. Stochastic influences

The optimisation model presented in Section 3.1 assumed that all model parameters are deterministic. In reality however, there are significant sources of uncertainty in offshore construction projects. Traditional stochastic project scheduling (see Section 2.2.1) defines probability distributions for each activity to express this risk.

For offshore construction projects such a model fails to capture the systematic uncertainty introduced by seasonal weather conditions. Hence, this research considers both uncertainty inherent to specific activities and the uncertainty introduced due to weather conditions.

The activity-dependent uncertainty is modelled by defining a triangular distribution for each of the activity durations, as is conventional in project scheduling literature. For the weather uncertainty, a novel simulation model is proposed based on recent advances in weather simulation techniques. The details of this weather simulation model are explained in detail in Section 4.2.1.

The implications for the scheduling optimisation model described in Section 3.1 are that both the activity durations ($d_i$) and the weather conditions ($\omega_{it}$) are now stochastic variables. Hence, the objective is not to optimise the gate times *NPV* for a single value of these parameters, but rather to optimise the expected value of the net present value ($E[NPV]$).

### 3.3. Example problem

This section presents a highly simplified example in order to illustrate the basic workings of the scheduling model. Assume a simple project consisting of two activities: 1 and 2, which are scheduled in series. The basic properties of these activities can be summarised in Table 2.

Both these activities depend on the same installation vessel for their execution. This vessel has a charter cost ($\rho_r^V$) of € 10 per period, and a commissioning ($\rho_r^C$) and decommissioning ($\rho_r^D$) cost of € 15. For this simple example it is assumed that there are no minimal periods required for chartering or vessel decommissioning.

The actual duration of both these activities also depends on the weather conditions. For this theoretical example, a simplified weather model which consists of only two states – good weather and bad weather – is used. Both activities are only capable of operating in the good weather conditions and are interrupted if weather conditions are bad. These weather states can be simulated using Markov chains defined by two transition matrices: one for the 'summer' season, and one for the 'winter' season. Furthermore it is assumed that each of these seasons lasts 15 time periods.

The actual value of the project itself (i.e. the cash inflow received upon project completion $\pi$) is equal to € 1250, and the discount rate per period (*ROI*) is equal to 1%.

Based on this information two basic strategies are immediately apparent. The first option is to start all activities as soon as possible. From the point of view of the decision variables this means that the first gate is set at time 1 (i.e. activity 1 can start immediately) and the second gate is set at the earliest possible starting time of activity 2. Assuming that activity 1 has the shortest possible duration of 9 time periods and the weather is continually favourable. This solution can be noted as: $(g_1, g_2) = (1, 10)$.

However, this strategy is likely to result in considerable delays because the second activity is executed at least partly in the winter period. Even in the best case scenario where there is not a single period of bad weather during summer and all activity durations are minimal. Hence, an alternative strategy is to delay the start of the second activity until the end of the winter period (i.e. $g_2 = 31$). This is likely to reduce the net working time during the project, but does delay the project's completion and results in a commissioning and decommissioning charge to be incurred, inflating the non-discounted cost of the project by €30.

Analysing the impact of these strategies requires simulating the execution of the project. A visual example of a single simulation run of the example problem is shown in Fig. 2. The simulation procedure has generated a specific duration for the activities from the uniform distributions in the table above: activity 1 has a duration of 10, and activity 2 has a duration of 11. Moreover, the simulation procedure has also simulated a weather pattern which is represented by the light and dark coloured blocks at the bottom of Fig. 2. A dark coloured block represents a period with bad weather, during which activity progress is interrupted. These interruptions extend the duration of the activities, as shown by the additional duration added at the end of activities for both schedules. An important note here is that making up for delays during the activity progress can only be done when weather conditions are good. Hence, bad weather conditions after the planned finish can further extend the duration of the activity as well – as can be seen from the duration of the second activity when using the first schedule.

The simulation shown in Fig. 2 indicates that weather conditions cause significant inflation of the duration of the activities. The effective activity durations are 12 and 18 for the first and the second activity respectively when the first schedule is used, and both activities have a duration of 12 when the second scheduling procedure is used (Table 3).

The scheduling strategies are represented visually by the 'gates' that indicate the earliest starting times of the activities. When comparing the performance of both schedules, it can be seen that the second schedule has set the gate time for the second activity much further, at time when good weather conditions were expected. This has resulted in a much shorter duration for the second activity, at the cost of inflating the duration of the complete project. To judge which of these schedules performed best during this simulation, the net present value of the costs and return of the project have to be analysed in detail, as is shown in Tables 4 and 5 which show the different components of the net present value as calculated using Eq. (1).

When comparing the results in detail it can be seen that the first schedule strategy results in a considerably higher variable cost for the second activity, as well as a greater chartering cost, since the total duration during which the vessels are used is longer (30 time periods). The second schedule on the other hand is able to present much lower costs, but decreases the value of the project by extending its duration, resulting in a lower net present value of the terminal value of the project.

Naturally, it must not be assumed that the results from a single simulation run are representative and more simulation runs are needed in order to get accurate estimates of the relative performance of the different scheduling strategies for this project.

## 4. Solution approach

This section presents the proposed solution approach for the problem defined in Section 3. Finding a solution for this problem is a complex problem because of two aspects. First and foremost the number of possible schedules is very large, as will be discussed in Section 4.1 where the solution space is discussed. Secondly, the stochastic nature of the problem requires simulations in order to correctly evaluate the quality of the schedule (see Section 4.2).

Fig. 3 gives a schematic overview of the proposed solution approach, including references to the sections in which the various elements are discussed. The two main building blocks of the solution procedure are the optimisation heuristics (Section 4.4) and the simulation procedures (Section 4.2). The optimisation heuristics create schedules for the construction project, which are expressed as different types of chromosomes depending on the optimisation heuristic used. These chromosomes are then passed to the simulation procedure which calculates the expected net

**Table 3**
Example weather model.

| | Good | Bad |
|---|---|---|
| Summer transition matrix | | |
| Good | 0.9 | 0.1 |
| Bad | 0.7 | 0.3 |
| Winter transition matrix | | |
| Good | 0.6 | 0.4 |
| Bad | 0.5 | 0.5 |

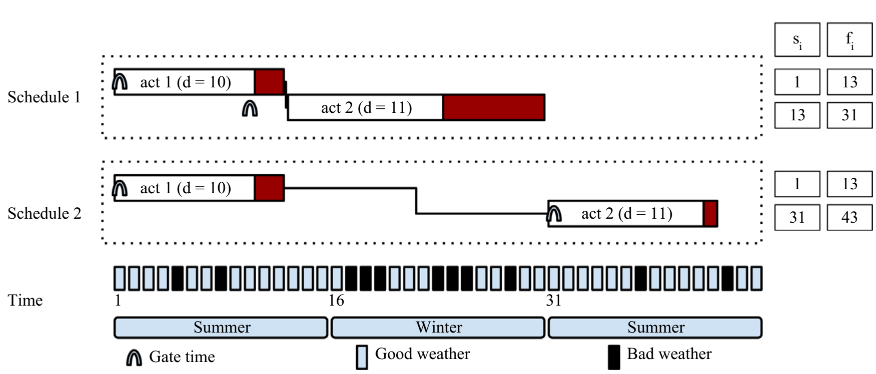


**Fig. 2.** Comparison of scheduling approaches for the sample problem.

present value ($E[NPV]$). This information is then returned to the optimisation heuristic.

Three different optimisation heuristics are proposed in this research: two meta-heuristic approaches (Sections 4.4.1 and 4.4.2) and a dedicated heuristic (Section 4.4.3). The chromosome encodings used by these algorithms have been created to reduce the size of the solution space, and by extension the performance of the solution heuristic, as discussed at length in Section 4.1. The simulation model uses discrete event based simulation (Section 4.2.2), and includes both stochastic activity durations and realistically generated weather patterns (Section 4.2.1). The weather model itself generates realistically correlated wind and wave states for every time period.

**Table 4**
NPV calculation for first scheduling strategy.

| Cost/proceeding component | Cashflow | Timing | NPV (€) |
|---|---|---|---|
| Activity 1 fixed cost | −100 | 0 | −100.0 |
| Activity 1 variable cost | −60 | 12 | −53.2 |
| Activity 2 fixed cost | −175 | 10 | −158.4 |
| Activity 2 variable cost | −180 | 30 | −133.5 |
| Chartering cost | −300 | 15 | −258.4 |
| Mobilisation | −15 | 0 | −15.0 |
| Demobilisation | −15 | 30 | −11.1 |
| Terminal value | 1250 | 30 | 927.4 |
| | | Total | 197.7 |

**Table 5**
NPV calculation for second scheduling strategy.

| Cost/proceeding component | Cashflow | Timing | NPV (€) |
|---|---|---|---|
| Activity 1 fixed cost | −100 | 0 | −100.0 |
| Activity 1 variable cost | −60 | 12 | −53.2 |
| Activity 2 fixed cost | −175 | 29 | −131.1 |
| Activity 2 variable cost | −120 | 41 | −79.8 |
| Chartering cost | −240 | 15 | −206.7 |
| Mobilisation 1 | −15 | 0 | −15.0 |
| Mobilisation 2 | −15 | 29 | −11.2 |
| Demobilisation 1 | −15 | 12 | −13.3 |
| Demobilisation 2 | −15 | 41 | −10.0 |
| Terminal value | 1250 | 41 | 831.3 |
| | | Total | 210.8 |

### 4.1. Solution representation

As was stated in Section 3.1, the operational schedule is defined by a set of gate times ($g_i$) for each activity $i$. This variable can have any positive integer value within the time interval taken into consideration $(1, \ldots, NT)$ and represents the moment in time when the materials needed for the execution of the activity have been made available. The model assumes that the payment for these materials has to be made at the moment they are delivered, meaning that the net present value of these outgoing cashflows is based on the gate times which have been set for these materials.

Based on this principle, several ways of constructing a solution representation (chromosome) can be proposed. When constructing such a solution representation it is important to balance the coverage of the total solution space with the total possible number of solutions. An encoding which is too detailed will increase the computation time needed to adequately search the solution space beyond practical value. A solution which is too restrictive on the other hand may cause severely suboptimal solutions due to the limited set of solutions which is considered. Sections 4.1.1–4.1.3 present three different solution encodings at different points in this spectrum. These encodings are then used by the heuristic solution procedures presented in Section 4.4.

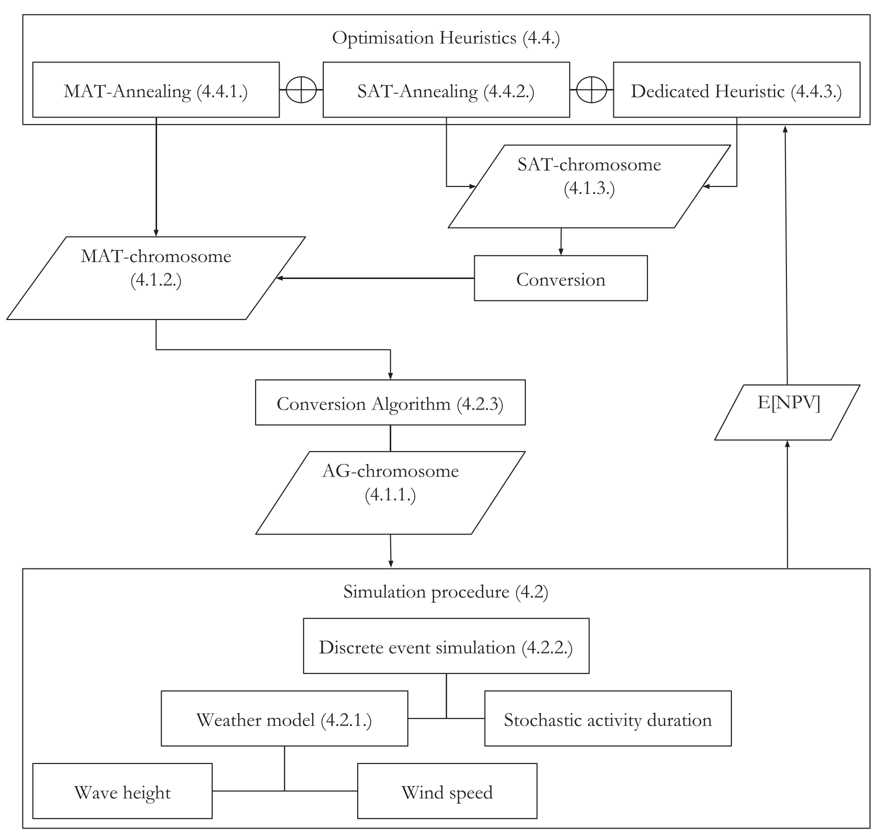


**Fig. 3.** Schematic overview of the solution approach.

*4.1.1. Activity gate chromosome*

The simplest way of representing a solution is simply to use the gate times for the individual activities as alleles in the chromosome encoding. Hence, each allele ($g_i$) is an integer gate time, indicating the specific time period in which the resources for activity $i$ are made available:

$$[g_1, g_2, \ldots, g_{NI}], \quad g_i \in \{1, \ldots, NT\} \tag{11}$$

The first schedule of the example presented in Section 3.3 can be represented by the following activity gate chromosome: $[1, 10]$. This implies that the resources for the first activity are available immediately ($g_1 = 1$), and the resources for the second activity are available at the start of the 10th time period ($g_2 = 10$). Practically this implies that the first activity will be able to start immediately, and the second activity will start at the 10th time period, or once the first activity has been completed, whichever is later: $s_2 = \max(g_2, f_1)$.

The disadvantage of this solution representation is the large size of the solution space, even for problems of limited size. The exact number of possible solutions can be calculated as $NT^{NI}$, where $NI$ is the total number of activities and $NT$ is the total number of time periods considered. Depending on the precedence relations in the project and the manner in which the durations of the activities are specified, a subset of these combinations can be eliminated from the pool of rational solutions. This is done by preventing gate times from taking values smaller than the earliest possible start of the respective activity.

*4.1.2. Monthly activity type chromosome*

Besides eliminating irrational solutions, the size of the solution space can also be reduced by using other chromosome definitions. One way of doing this is by taking advantage of the repetitive nature of offshore projects: many types of activities, and even sequences of activities are repeated within the same project. Hence, rather than specifying when specific activities are allowed to start, it is possible to define during which time periods activity types are allowed to be executed.

The size of the solution space can be reduced even further by aggregating the time periods into longer intervals. Given that the weather model uses months as climatological units, and the motivation for allowing or not allowing certain activity types is likely motivated by expected weather conditions, it makes sense to aggregate the time periods per month.

The term 'monthly activity type' or *MAT* chromosome will be used to describe this type of chromosome. The chromosome itself consists of a set of binary variables $x_{j\tau}$, which signify whether (1) or not (0) activity type $j$ ($j \in [1, 2, \ldots, NJ]$) is allowed to be executed in month ($\tau\tau \in [1, 2, \ldots, N\tau]$). The chromosome can be visualised as a matrix where the rows represent the activity numbers and the columns represent the time periods (months):

$$\begin{bmatrix} x_{11} & \ldots & x_{1,N\tau} \\ \ldots & x_{jm} & \ldots \\ x_{NJ,1} & \ldots & x_{NJ,N\tau} \end{bmatrix} \tag{12}$$

Again this chromosome can be illustrated using the simple example which has been defined in Section 3.3. For this example the 'summer' and 'winter' seasons can be seen as the relevant time periods indexed $\tau = 1, 2, 3$. The two activities in the example project are assumed to be of different types: $j = 1, 2$. This results in a total of six binary variables which can be placed in a matrix structure to represent the *MAT* chromosome encoding:

$$\begin{bmatrix} x_{11} = 1 & x_{12} = 0 & x_{13} = 0 \\ x_{21} = 0 & x_{22} = 0 & x_{23} = 1 \end{bmatrix} \tag{13}$$

It can be noted that the values of $x_{12}$ and $x_{13}$ could also be set to 1 without impacting the way in which the project is executed. This is not a violation of the 1 to 1 equivalence of a chromosome and a solution to the problem. The underlying execution strategies for both these chromosomes are indeed different. However these differences held no relevance with regard to the outcome of the project. Hence, the fact that these variables can take different values should simply be seen as an ex aequo between two different solutions.

The substantial reduction of the solution space when using the *MAT* chromosome encoding, rather than the gate times encoding is already apparent from this small example. Using the *MAT* encoding, the total number of combinations can be calculated as $2^{NJ \cdot N\tau}$, for the example from Section 3.3 this means that there are a total of $2^{2 \cdot 3} = 64$ possible combinations. Whereas when using the original chromosome the total number of combinations was equal to $NT^{NI} = 45^2 = 2025$.

In order to determine the expected net present value when using a schedule defined by a *MAT*-chromosome the activity specific gate times have to be determined (see Section 3.1). Effectively, the higher-level *MAT*-chromosome has to be translated into an *AG*-chromosome (see Section 4.1.1) encoding. This is done using an iterative simulation procedure which is explained in detail in Section 4.2.3.

*4.1.3. Seasonal activity type chromosome*

A further reduction of the solution space can be obtained by reducing the project schedule to a combination of two key decisions. The first decision being when to start the project, and the second decision being how to deal with the weather sensitivity for each activity type.

Similar to the *MAT*-chromosome encoding, this chromosome uses the climatological periods as they are defined in the weather simulation model (months for the model used in this research). However, whereas the *MAT*-chromosome defined a variable for each occurrence of these time periods (i.e. January of 2015, January of 2016, etc.), the encoding scheme only defines a single variable for each climatological period (i.e. January, February, etc.). In effect, the seasonality of the weather pattern is used to decrease the complexity of the scheduling decision. Hence, this type of chromosome encoding is dubbed a seasonal activity type (*SAT*) chromosome.

The *SAT*-chromosome is a list of integer variables in the range $[0, NP-1]$ where $NP$ is equal to the number of climatological periods used in the weather model, in this case 12 months. This list includes one variable $s$ which indicates the number of months the start of the project is delayed. This allele can be used to align periodic weather sensitivity of the project with periods of more favourable weather. A value of $s=0$ indicates that the project starts immediately, whereas a value of $s=3$ indicates that the start of the project is postponed with three months (i.e. assuming the earliest start would be in January, a value $s=3$ will result in the project starting in April).

The second part of the *SAT*-chromosome consists of a variable ($d_j$) for each of the activity types ($j$). This variable represents the number of months during which activities of type $j$ will not be performed to avoid weather related delays. Naturally, the months during which the execution is halted are the months during which the worst average weather conditions are expected. Which time periods are the worst on average is determined using the steady state probability matrices of the weather model (see Section 4.2.1). The relative weather quality is simply defined as the average of the wind- and wavestates during the time period.

This results in the following representation of the project schedule:

$$[s, d_1, d_2, \ldots, d_j, \ldots, d_{NJ}], \quad s, d_j \in [0, \ldots, NP-1] \tag{14}$$

By using this encoding, the number of possible schedules in the solution space is further reduced to $(1+NJ)^{NP}$. Again the use of this chromosome can be illustrated using the example from Section 3.3. Again assuming that the two activities in the example are of different types, the following *SAT* encoding can be used to represent the second schedule of the example:

$$[s=0, d_1=0, d_2=1] \tag{15}$$

Eq. (15) shows that the schedule starts immediately during the first period ($s=0$). Also the first activity is allowed to be executed during every time period, whereas the second activity cannot be executed during the period with the worst weather conditions. Which period has the best or worst weather or average is determined by using the steady state probabilities, as described in Section 4.2.1. For the simple example the steady state probability for good weather during the summer season is 87.8%, whereas the probability for good weather is equal to only 55.6% during the winter season. Hence, it is clear that the winter season has the worst weather conditions and the second activity type should be suspended during the winter rater than the summer season.

In order to obtain the specific gate times associated with the schedule represented by the *SAT*-chromosome, the *SAT*-encoding is simply translated into a *MAT*-chromosome, after which the procedures used for the latter can be used to determine the activity specific gate times.

#### *4.1.4. Comparison of solution space size*

The reduction of the size of the solution space is illustrated in Fig. 4. This example assumes a project with a five-year timespan using 4-h time periods for the simulation procedure. It is clear that the chromosome used has an impact of several orders of magnitude regarding the size of the solution space. Moreover, this impact increases as the size of the project increases.

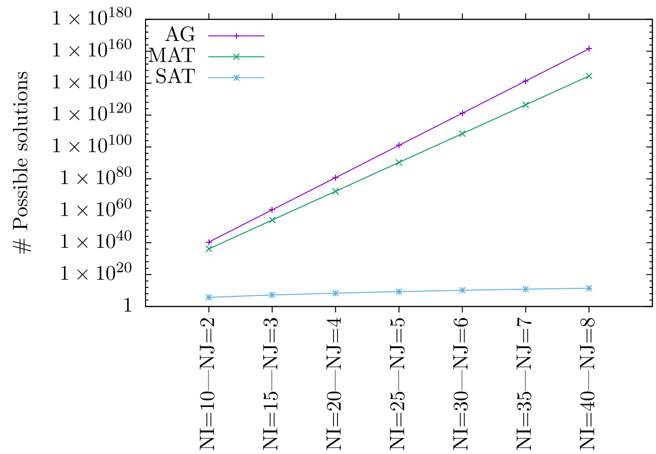


**Fig. 4.** Evolution of the size of the solution space when using different chromosomes.

### *4.2. Simulation procedure*

#### *4.2.1. Modelling the weather*

Using historic wind speed and significant wave height data collected off the Belgian coast in the North Sea,[2] a simulation model was created. This model uses a combination of transition probability matrices and Weibull distributions to replicate a realistic combination of wind and wave conditions faced by offshore construction projects in the North sea.

The main difficulty encountered when creating this model was simultaneously guaranteeing the right correlation of wind speed and wave height, as well as the individual autocorrelation of both these variates. This was achieved by defining conditional probabilities that link these variables to each other as well as to past observations.

Because the main objective of this simulation is to determine the economical impact of the weather conditions, rather than a detailed description of meteorological phenomena, both wind and wave data have been transformed into discrete states. The evolution of these weather states is described using 4-h intervals. This interval is somewhat longer than that of models with meteorological objectives. The reason for this is dual: firstly, very short time intervals are not relevant from an economic perspective. Secondly, using longer time intervals allows for a better approximation of autocorrelation.

Fig. 5 shows the main flow of the weather simulation algorithm. Because substantially more wind data was available,[3] the starting point is the generation of the wind speeds. These wind speeds then form the basis to generate associated wave heights.

*Wind speed*

The simulation of wind states is done using separate first-order Markov chains for every month of the year, to account for the inherent non-stationarity of the data. Each of these models uses a transition matrix which defines a probability $p_{ij}^{wind}$, signifying the likelihood of moving from wind state $i$ to wind state $j$ during a certain season. Naturally these probabilities have to satisfy the following property:

$$\sum_{j=1}^{n} p_{ij}^{wind} = 1, \quad \forall i \tag{16}$$

where $n$ signifies the total number of possible wind states. This property simply states that for every wind state $i$, the sum of all the transition probabilities must be exactly equal to 1. The exact value for these probabilities ($p_{ij}^{wind}$) has been calculated for every month based on the available weather data.

Given that these Markov models for the wind states are ergodic (i.e. any state can be reached from any other state in a finite number of steps), it is possible to calculate the equilibrium probabilities by means of simulation. These simulations start at a random point and perform a certain number of iterations ($\iota$), after which the frequencies at which specific states were visited are transformed into probabilities. After the probabilities have been updated the steady state estimates are compared to the previous estimates, if the change exceeds a preset threshold the simulation continues for another $\iota$ iterations. For the model presented in this paper a threshold of 0.1% and an $\iota=100$ was used. The steady state probabilities obtained using this method are used to generate a starting point for the weather generation model (see Fig. 5).

A trade-off has to be made between the amount of information conveyed by the simulation by increasing the number of wind states, and the accuracy of the model, which decreases as the number of wind states increases because a larger number of parameters has to be estimated. The calibrations of the simulation model have indicated that the accuracy of the model starts to

[2] This data was collected by the Measurement Network Flemish Banks, specifically the data from the "Westhinder" measuring pile was used since this location most accurately reflects weather conditions for offshore wind farms off the Belgian coast. More information can be found at: http://www.vliz.be/en/measurement-network-flemish-banks

[3] 4897 days of wind data and 1396 days of wave data were available at the time of this research.

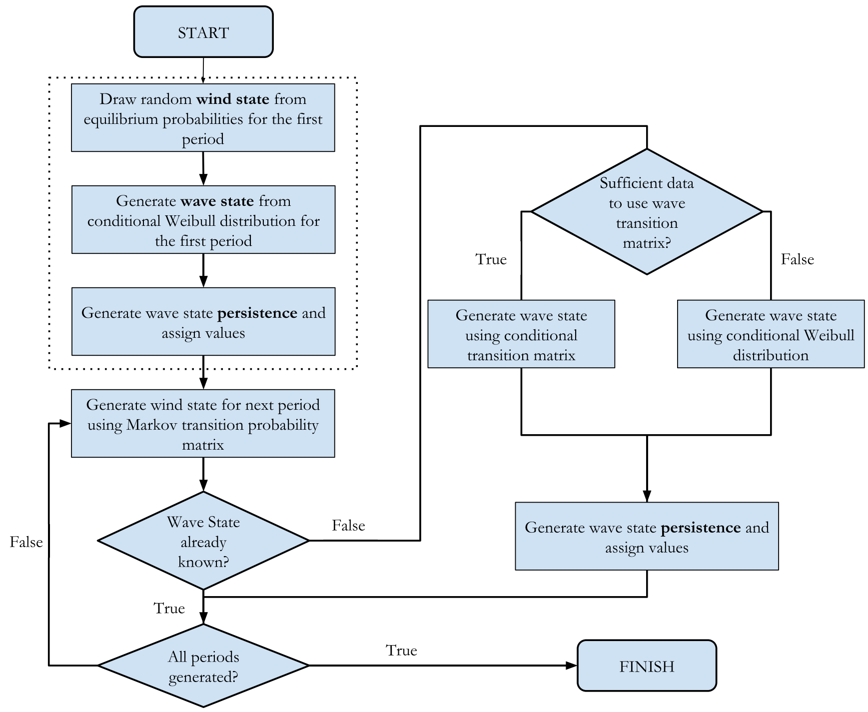


**Fig. 5.** Flowchart of the weather generation model.

decrease when more than 10 wind states are used. Hence, the total number of wind states has been set to 10 ($n = 10$). Note that the bounds for these states are simply the $n$th percentiles of the available data.

*Significant wave height*

The simulation of realistic wave heights to accompany the wind speeds is more complex for a number of reasons: (i) the wave state has to be correlated with the wind state, (ii) the wave state also has to have the right level of autocorrelation, (iii) the persistence of a wave state has to be realistic, and (iv) the limited amount of data points – especially for rare weather conditions – for fitting have to be taken into account.

To accommodate all these aspects, a combination of three techniques is used: two of which simulate realistic transitions between different wave states, and a third which simulates the persistence of a wave state once it is attained. The difference between the two techniques which simulate transitions is that the first one takes into account both the correlation with the wind state and the autocorrelation of the wave states, whereas the second technique only takes into account the correlation with the wind state. Which of these techniques is used is a response to the amount of data which is available to estimate the transition probabilities; in case insufficient data is available to use the first model the second model is used.

The first model which simulates wave state transitions uses a variation on a second-order Markov chain. This variation defines the possibility of transitioning to another wave state in period $t$ based on the preceding wave state from period $t-1$, as well as the already generated wind state for period $t$. Hence, a three-dimensional matrix containing transition probabilities $p_{kli}^{wave}$ is created. The value of $p_{kli}^{wave}$ represents the probability of moving from wave state $k$ in period $t$ to wave state $l$ in period $t+1$, given that the wind state (which has already been simulated using the wind state Markov chain) in period $t+1$ is equal to $i$. Again, for every possible value of $k$ and $i$, the sum over all possible $l$ values ($l = 1, \ldots, m$) has to equal unity:

$$\sum_{l=1}^{m} p_{kli}^{wave} = 1, \quad \forall k, i \tag{17}$$

Because the persistence of these wave states is modelled separately, the probability of moving to the same wave state is set to zero. Using the same notation this implies that:

$$p_{kki}^{wave} = 0, \quad \forall k, i \tag{18}$$

By using this transition matrix, the correlation between the wind and wave data, as well as the autocorrelation of the wave data can be guaranteed. However, the drawback of this method is that it requires a lot ($n^2 \cdot m$) of probabilities to be estimated. Especially for rare scenarios where winds are relatively high and waves relatively are low or vice versa it may not be possible to estimate realistic probabilities by simply fitting them to the available data because no such observations exist. Nevertheless, the simulation model should be capable of dealing with these unlikely scenarios. This is the reason for the introduction of the second method for generating correlated wave states.

The second method for simulating transitions between wave states only takes the correlation between the wind and wave states into account. Depending on the wind state which has already been simulated, the accompanying wave state is drawn from one of $n$ Weibull distributions. These $n$ distributions have been fitted to subsets of the wave data, based on the wind state which accompanied the specific wave state. Assume for example that a wave state has to be simulated, given the fact that the wind state for the specific period has already been simulated and is equal to 2. Moreover, there is insufficient data to accurately estimate parameters using the first model. Hence, the wave state will be randomly drawn from a Weibull distribution whose parameters have been fitted to the wave heights which occurred when the wind state was also equal to 2. Note that this Weibull model is both fitted to and simulates continuous data rather than the discrete wave states, requiring the wave state to be determined after a continuous Weibull distributed random variable has been returned.

Similar to the first method of generating the wave state transitions, the Weibull model also does not allow the wave state to remain unchanged. Should an identical wave state be drawn from the distribution, the draw is discarded and a new Weibull distributed random number is drawn.

The persistence of good and bad weather conditions is one of the most important aspects of a weather simulation model [7,16]. Hence, one of the major goals of this situation model was to give an accurate representation of the persistence of weather conditions. Preliminary experiments which used the previously described models to simulate both transitions to different and the same wave state concluded that these techniques could not guarantee a realistic persistence of the wave states. This is caused by the correlation which is introduced between the wind and wave states.

This issue was solved by generating the persistence of a wave state separately. This is done by creating a persistence probability matrix for every wave state based on the available wave data. This matrix consists of probabilities $p_{kt}^{persist}$, which signify the probability that wave state $k$ persists for $t$ time periods. Naturally, the following must be true:

$$\sum_{t=1}^{t_{\max}} p_{kt}^{persist} = 1 \quad \forall k \tag{19}$$

where $t^{\max}$ represents the maximal number of time periods a wave state can persist. Using these conditional probabilities, the persistence of every wave state is modelled. Analysis shows that by using this method of simulation satisfactory results are obtained for the autocorrelation of both wind and wave states (see Appendix A).

The cost of this improved autocorrelation of the wave states is a slight decrease of the correlation between wind and wave states. In the original data a correlation of 0.8 was noted between wind and wave states. Due to the fact that the persistence of a wave state is assumed to be independent from the wind states, the global correlation between these two variates is reduced to 0.7. Insufficient data inhibits the use of a correction of the Weibull distribution without overfitting the available data points. Artificially inflating the correlation of the wind and wave states on the other hand results in a substantial reduction of the frequency of the wave states.

During the calibration of the simulation algorithm it was concluded that setting the number of available wave states to 10 ($m=10$) provided the best compromise between model complexity and accuracy in representing the weather patterns.

#### 4.2.2. Simulation of AG-chromosome solutions

The goal of the simulation procedure is to obtain a value for the start and finish times for every activity based on a *AG*-chromosome schedule representation (see Section 4.1.1), hereby allowing the net present value (*NPV*) of the project to be calculated (see Section 3.1). The simulation procedure itself takes into account two distinct sources of uncertainty: uncertainty with respect to the duration of activities and uncertainty with respect to the influence of weather conditions.

**Algorithm 1.** Simulation procedure.

1 : **procedure** RunSimulation (*AG*-chromosome, weatherPattern)<br>
2 :   agenda $\leftarrow \varnothing$<br>
3 :   $\forall i$ : agenda $\leftarrow$ agenda $\cup \{g_i\}$<br>
4 :   $\forall i$ : $released_i \leftarrow False$<br>
5 :   $\forall i : \rho_i \leftarrow$ # predecessors$_i$<br>
6 :   **while** agenda $\neq \varnothing$ **do**<br>
7 :     $e \leftarrow$ agenda.pop()<br>
8 :     $\iota \leftarrow e.index()$<br>
9 :     **if** e.type() = gate **then**<br>
10 :       $released_\iota \leftarrow True$<br>
11 :       **if** $\rho_\iota = 0$ **then**<br>
12 :         $s_\iota \leftarrow$ e.time()<br>
13 :         agenda $\leftarrow$ agenda $\cup \{f_\iota\}$<br>
14 :     **else if** e.type() = finish **then then**<br>
15 :       **for** s in successors$_\iota$ **do**<br>
16 :         $\rho_s \leftarrow \rho_s - 1$<br>
17 :         **if** $\rho_s = 0$ and released$_s$ $= True$ **then**<br>
18 :           $s_\iota \leftarrow$ e.time()<br>
19 :           agenda $\leftarrow$ agenda $\cup \{f_s\}$<br>
20 :     **return** $\{s_i, f_i\} \quad \forall i$

Algorithm 1 is a pseudocode representation of the discrete-event based simulation procedure used to evaluate the performance of a specific schedule. The input for the simulation procedure is an *AG*-chromosome. Such a chromosome consists of a gate time $g_i$ for each of the activities in the project. The outcome of the simulation procedure is a starting ($s_i$) and finishing time ($f_i$) for each of the activities. This information is sufficient to calculate the net present value for this specific scenario. Naturally in order to obtain an estimate of the expected net present value ($E[NPV]$) multiple simulations will have to be performed.

The simulation procedure starts by defining an empty event agenda, and then adds the gate times ($g_i$) for all the activities as events to the agenda (line 3). For each activity a boolean variable $released_i$ is defined which is used to indicate if the gate time of activity $i$ has already occurred. Another variable $\rho_i$ is introduced and initialised to track the number of unfinished predecessors remaining for every activity.

Once these variables have been declared, the procedure starts looping over the events in the agenda (line 6). At each iteration the next event $e$ is removed from the event agenda. The variable $\iota$ is used to represent the index of the activity to which the event is linked. Two types of events are considered in the simulation procedure: a gate time of an activity and the finishing time of an activity.

If the event $e$ is a gate time of an activity (line 9), the $released_\iota$ variable is set to *True*. This indicates that the activity can start as soon as all its predecessors have been completed. Next, the procedure tests if there are no remaining predecessors (line 11). If this is the case, the finishing time of the activity can be calculated, taking into account the a randomly drawn activity duration, as well as the simulated weather conditions in combination with the operational restrictions of the specific activity. This finishing time is then added to the agenda as a new event since it may indicate the time at which the number of predecessors of an activity has to be decreased.

If the event $e$ is a finishing time of an activity (line 14), the algorithm iterates over all the successors of activity $\iota$, decrementing the number of remaining predecessors ($\rho_s$) for each of these successors. For each successor, the procedure also tests if the total number of predecessors is now zero. If this is the case and the activity has already been released the starting and finishing times of the activity can be determined. The finishing time of the activity is then also added as a new event to the agenda.

#### 4.2.3. Chromosome re-encoding simulation

As was mentioned in Section 4.1.2, the higher-level chromosome encodings (*MAT* and *SAT*) do not directly specify gate times for all individual activities. Nevertheless, simulating the execution of the project using the procedure described in Section 4.2.2 requires explicit gate times to be defined for each activity. Naturally, due to the stochastic nature of the project it is infeasible to specify gate times which are guaranteed to respect the conditions

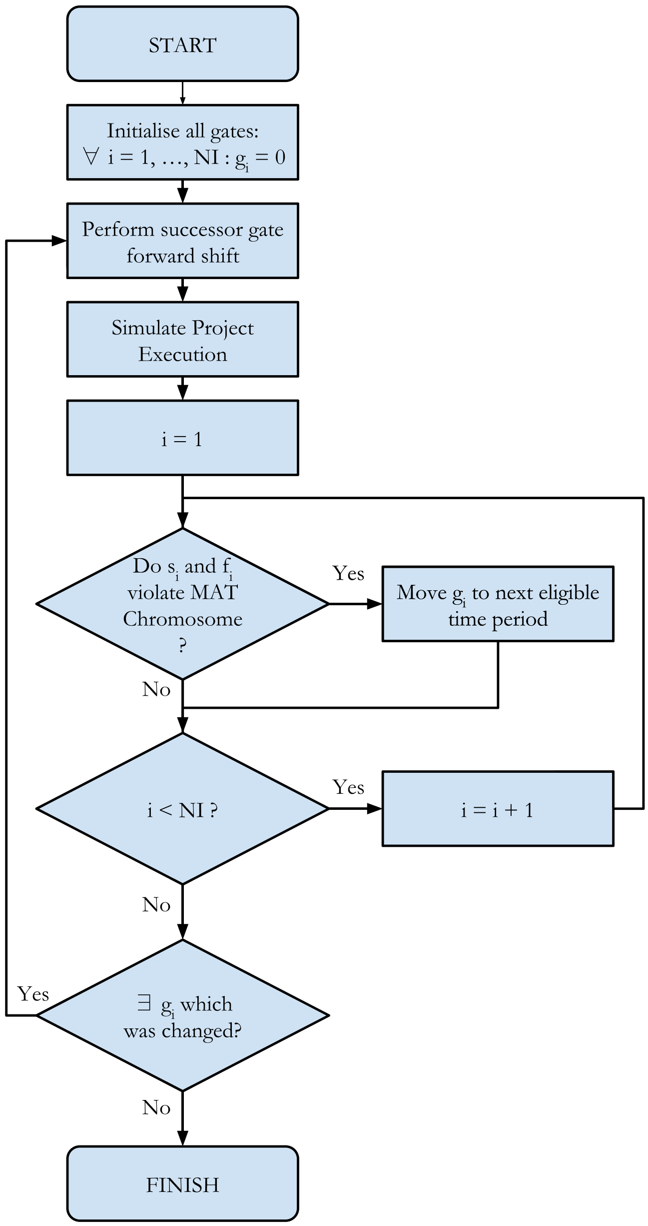


**Fig. 6.** Iterative procedure to set gate times.

set by the higher order *MAT* or *SAT* chromosomes. Because of this, an iterative satisfying procedure to set the gate times is proposed. The outline of this procedure is visualised by Fig. 6.

The procedure starts by setting all gate times ($g_i$) at the start of the project ($g_i = 1$). Next, the precedence relations are used to shift a fraction of these gate times forward in time based on the release times and shortest possible durations of their predecessors (assuming no weather delays). This set of gate times is then simulated using the procedure which was described in detail in Section 4.2.2. The outcome of this simulation is a set of starting ($s_i$) and finishing times ($f_i$), which are compared with the conditions imposed by the *MAT* or *SAT* chromosome. In case the execution of an activity falls within a period which does not allow execution of its activity type, the gate time of the activity is shifted forwards.

Once this has been tested for all the activities, the algorithm verifies if any gates were shifted since the last simulation run. If this is the case, the forward shift procedure for all successors is called again, and a new simulation is performed. When no more gate times have been adjusted after performing a simulation, the restrictions imposed by the higher-order chromosome are assumed to be satisfied.

### 4.3. Comparing solutions

Due to the stochastic nature of the scheduling problem, it is important to verify that differences in performance between two schedules are statistically significant. This is especially important when comparing alternatives in a (meta-)heuristic procedure. However, constructing independent confidence intervals for both solutions is computationally expensive. When using unrelated samples the variance of the difference between two estimators can be calculated as:

$$\text{var}\left[N\hat{P}V(s_1) - N\hat{P}V(s_2)\right] = \text{var}\left[N\hat{P}V(s_1)\right] + \text{var}\left[N\hat{P}V(s_2)\right] \tag{20}$$

where $N\hat{P}V$ is an estimator of the *NPV* for a specific schedule *s*. Naturally, the statistical significance of the difference between two solution can be improved by increasing the number of simulations. However, this increased statistical significance comes at a considerable computational expense.

Constructing correlated samples by using common random numbers is an alternative technique which can also increase statistical significance at no increased computational expense. By comparing the performance difference between two schedules using the same random estimates for the weather conditions and the activity durations, a strong correlation of the samples can be obtained. Whereas the covariance between the two estimators could be assumed to be zero in Eq. (20), given the independent samples used, the covariance will now become strictly positive. This means that the variance of the difference between the two estimates now becomes:

$$\text{var}\left[N\hat{P}V(s_1) - N\hat{P}V(s_2)\right] = \text{var}\left[N\hat{P}V(s_1)\right] + \text{var}\left[N\hat{P}V(s_2)\right] - 2 \cdot \text{cov}\left[N\hat{P}V(s_1), N\hat{P}V(s_2)\right] \tag{21}$$

Given that the covariance will always be strictly positive, this will always be a reduction of the variance of the difference between the two estimators. Hence, this technique enables a more efficient one-on-one comparison of different schedules, increasing the number of schedules which can be tested in a given timeframe. The latter being of paramount importance to the performance of local-search based meta-heuristic procedures since it allows a larger number of solutions to be investigated in a given timeframe.

### 4.4. Optimisation heuristics

Creating optimal schedules is complicated by both the size of the neighbourhood (see Section 4.1) and the need for simulation in order to estimate the *NPV* associated with a specific schedule. Especially since these simulations are relatively computationally expensive, enumerative procedures which iterate over all possible schedules are eliminated as viable solution methods. Moreover, exact solution techniques are also ill-suited due to the stochastic nature of the problem.

Hence, the scheduling optimisation uses heuristic techniques to make an efficient scan of the most promising areas of the solution space. In the remainder of this paper three different heuristics to solve this problem are proposed and tested. Two of these are local-search based meta-heuristics, the first of which uses the *MAT*-chromosome encoding, and the second of which uses the higher-level *SAT*-chromosome encoding. Note that the choice for local search rather than population based meta-heuristics was made because of the efficiency gains which are possible when comparing solutions one-on-one using common random numbers (see Section 4.3). These efficiency gains can no longer be obtained when selecting the best solutions out of a large pool of candidates.

Rather than using general purpose techniques on a specific solution encoding, the third solution technique is a dedicated heuristic which tests the most 'rational' choices in the solution spectrum. The subsequent sections provide more information on the specifics of the meta-heuristics which are used, after which the performance of these heuristics is tested in Sections 5 and 7.

*4.4.1. MAT-chromosome annealing*

The first meta-heuristic is an implementation of the simulated annealing meta-heuristic Talbi, [45], which uses the *MAT*-chromosome encoding presented in Section 4.1.2. A basic outline of the annealing procedure is presented by Algorithm 2.

**Algorithm 2.** Basic annealing procedure.

1 : **procedure** SIMULATED ANNEALING (problem instance)
2 : $t \leftarrow t_{\text{initial}}$
3 : $i \leftarrow 0$
4 : $\text{sol}_{\text{best}} \leftarrow$ random init
5 : $\text{sol}_{\text{incu}} \leftarrow \text{sol}_{\text{best}}$
6 : **while** $t > t_{\text{end}}$ and $i <$ maxiter **do**
7 : NeighMove($\text{sol}_{\text{incu}}$)
8 : $\delta = \text{significantDifference}(\text{sol}_{\text{best}}, \text{sol}_{\text{incu}})$
9 : **if** $\delta < 0$ **then**
10 : $\text{sol}_{\text{best}} \leftarrow \text{sol}_{\text{incu}}$
11 : $t \leftarrow t \cdot$ cooldown
12 : **else if** $\delta = 0$ **then**
13 : Insignificant difference:
14 : $\text{sol}_{\text{incu}} \leftarrow \text{sol}_{\text{best}}$
15 : **else if** $e^{-\frac{\delta}{t}} < r$ **then**
16 : $\text{sol}_{\text{best}} \leftarrow \text{sol}_{\text{incu}}$
17 : $t \leftarrow t \cdot$ cooldown
18 : **else**
19 : $\text{sol}_{\text{incu}} \leftarrow \text{sol}_{\text{best}}$
20 : $i \leftarrow i + 1$
21 : **return** $sol_{best}$

The neighbourhood moves are performed by flipping a random binary value in the *MAT*-chromosome from 1 to 0 or vice versa. Next, the iterative procedure (see Section 4.2.3) is used to create an *AG*-chromosome which can be used by the simulation procedure which estimates the *NPV* of the solution.

The meta-heuristic also verifies that infeasible solutions are discarded. This can occur when a certain solution does not allocate sufficient time to a certain activity type in order to complete it within the timeframe of the simulation procedure. In case such an infeasible solution is detected, the solution is discarded and a new neighbourhood move is tested.

As shown in Algorithm 2, the heuristic starts by initialising the temperature ($t$) and the iteration counter ($i$). The current best solution is initialised randomly, and the incumbent is set to be equal to this initial solution (lines 3 and 4). Next the iterations of the simulated annealing heuristic start, and continue until the temperature has become lower than the predefined end temperature ($t_{end}$) or the maximal number of iterations has been reached (line 5). During each iteration a random neighbourhood move is performed on the incumbent solution (line 6). Next, the difference between the best ($sol_{best}$) and incumbent solution ($sol_{incu}$) is analysed using the procedure described in Section 4.3. This comparison yields one of the four outcomes: (i) the solution is better and replaces the current best solution (line 9), or (ii) there is no significant difference between the current best and the incumbent and the move is discarded (line 12), or (iii) the solution is significantly worse but is still accepted according to a certain random probability expressed as a function of the current temperature (line 15), or (iv) none of the above is true and the solution is discarded (line 18). A non-traditional element of this simulated annealing meta-heuristic is that it discards any moves which do not yield a statistically significant difference between the current and the incumbent solution.

*4.4.2. SAT-chromosome annealing*

A variation on the simulated annealing procedure presented in Section 4.4.1 can be created by using the *SAT* chromosome rather than the *MAT* chromosome. The neighbourhood operator for this meta-heuristic increments or decrements a single allele of the chromosome encoding, naturally taking into account the bounds of the individual alleles ($s, d_j \in [0, 11]$). Similarly to the previous meta-heuristic, infeasible solutions due to too little available production time are discarded. All other elements of the heuristic remain unchanged.

*4.4.3. Dedicated heuristics*

For problems where evaluating a solution takes a considerable amount of computation time, dedicated heuristics can provide a valuable alternative. Rather than using a random approach to explore the solution space, dedicated heuristics use as much problem information as possible to explore the most promising regions. This section presents a dedicated heuristic which takes a subset of the solution space as defined by the *SAT*-chromosome and uses a full-factorial approach to explore all solutions therein. This subset can be defined as the *SAT*-chromosomes (see Eq. (14)) for which the following condition holds:

$$d_j = d_k \quad \forall k, j = 1, \ldots, NJ \,|\, j \neq k \tag{22}$$

Eq. (22) simply states that the number of months during which activity types are halted because of averse weather conditions is equal across all activity types. This effectively reduces the *SAT*-chromosome to the following expression:

$$[s, d], s, d \in [0, \ldots, NP - 1] \tag{23}$$

where $d$ represents the number of months of downtime for adverse weather conditions for all activities. Hence, the total number of combinations is reduced to $NP^2$. Given that the number of seasons distinguished by weather models rarely exceeds the 12 months of the year. Hence, the computational effort needed to calculate all these combinations is unlikely to exceed computational limitations, even for realistically sized projects.

## 5. Case study: the construction of an offshore wind farm

The objective of this research is to present tools which can be used for realistically sized problems which are subjected to substantial weather risks. An example with increasing relevance in recent years is the construction of offshore wind farms. This section presents a project consisting of 54 wind turbines which have been installed on the Thorntonbank off the Belgian coast, having a total capacity of 325 MW.

Given that the objective is to optimise the scheduling of the project taking into account the weather conditions, the activity breakdown structure should be sufficiently detailed to incorporate all differences in weather sensitivity between activities. Practically this means that 12 different activity types are distinguished. The properties of these activity types are listed in Table 6. The following paragraphs provide a more detailed insight in the properties of these activities, as well as their interrelations.

**Table 6**
Costs, durations and weather restrictions associated with the different activity types.

| Activity type | Activity subtype | Costs in 000 € | | $\Delta$-duration (h) | Weather restrictions | | | |
|---|---|---|---|---|---|---|---|---|
| | | $C_i^F$ | $C_i^V$ | | Wave (m) | Wind (m/s) | Wavest. | Windst. |
| Scour protection | Sour protection | 64.87 | 3.17 | (20, 24, 28) | 1.5 | – | 5 | – |
| Foundation | Loading | 3192.69 | 9.72 | (5, 6, 8) | – | 10 | – | 4 |
| | Transport | 0.00 | 9.72 | (4, 6, 8) | 2 | 14 | 7 | 7 |
| | Positioning | 0.00 | 9.72 | (2, 4, 5) | 1.75 | 10 | 6 | 4 |
| | Installation | 0.00 | 9.72 | (11, 15, 22) | 1.75 | 9 | 6 | 3 |
| | Return to port | 0.00 | 9.72 | (4, 6, 8) | 2 | 14 | 7 | 7 |
| Turbine | Loading | 8779.91 | 9.72 | (9, 10, 12) | – | 10 | – | 4 |
| | Transport | 0.0 0 | 9.72 | (4, 6, 8) | 2 | 14 | 7 | 7 |
| | Positioning | 0.00 | 9.72 | (2, 4, 5) | 1.75 | 10 | 6 | 4 |
| | Installation | 0.00 | 9.72 | (24, 27, 34) | 1.75 | 8 | 6 | 2 |
| | Return to port | 0.00 | 9.72 | (4, 6, 8) | 2 | 14 | 7 | 7 |
| Inner array cable | Inner array cable | 508.84 | 28.50 | (40, 80, 160) | 1.75 | – | 6 | – |

In order to construct an offshore wind farm, four basic tasks must be performed. First and foremost, the place where the foundation will be placed has to be prepared by installing scour protection ($S$). This is done by using a stone dumping vessel to place rocks that prevent ocean current from digging out the monopile foundations out from under the wind turbine. Secondly, the foundation of the turbine ($F$) has to be installed using a jack-up barge which hammers the foundations into the seabed. Once this has been completed, the third activity is installing the turbine ($T$) itself, including the tower, navette and blades. Finally the fourth task is to dig a cable ($C$) into the seabed which connects each of the turbines to the transformer station. This final installation is done using a specially equipped trenching vessel.

Fig. 7 presents a high-level overview of the interrelations between the different activities. Naturally, the order which has been described has to be followed for each individual turbine: first the scour protection is installed, then the foundation of the turbine is installed, next the turbine itself is placed on top of the foundation and finally the cables are installed. However, because of the technical capabilities of the vessels used for these installations the installing of the foundation and the placement of the turbine are grouped. This because the jack-up barge is capable of transporting three foundations or three turbines simultaneously. Hence the pattern as is shown by Fig. 7 is created.

A more detailed view is shown by Table 6, which shows that both the installation of the foundation and turbine can be split up into more detailed sub-activities: loading the barge with the materials (incurring a substantial fixed material cost $C_i^F$), transporting the materials to the installation site, positioning of the barge for the installation, the installation of the turbine or foundation and the return to shore. There is also an operating cost associated with each of these activities ($C_i^V$), which represents the variable operating cost of the vessel and its crew, this is the surplus cost compared to the costs which are incurred when a chartered vessel is idle. The chain of events also respects the order which is shown by 7: first materials for three foundations or three turbines are loaded, after which these are transported to the site. Next the barge needs to be positioned for each turbine location after which the associated installation takes place. When all materials which were transported have been installed the barge returns to the dock for its next shipment.

Each of these activities has an associated triangular distribution[4] which represents the duration of the associated activity. Table 6 also shows the weather restrictions for each of the activities. It can be seen that the installation of the turbine is most susceptible to averse weather conditions, followed by the installation of the foundation, placement of the inner-array cable and the scour protection respectively.

[4] These triangular distributions are defined as: (min, modulo, max).

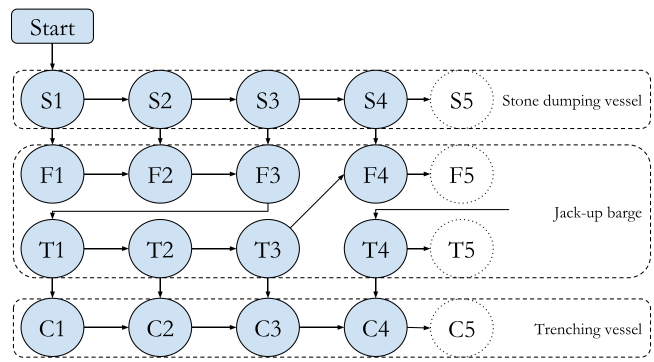


**Fig. 7.** Wind farm project structure.

The scheduling procedure as defined in the preceding sections is implemented on the level of the activity sub-types. The reason for this being the more realistic representation of weather sensitivity as well as the uncertainty related to each of the sub-types. Hence, 12 activity types are distinguished in the scheduling algorithms used to solve this case study.

Table 7 provides more information on the properties of the vessels used to execute the activities. As indicated in Fig. 7, the scour protection is installed by a stone dump vessel, the installation of both the foundation and the turbine is done by a general-purpose jack-up barge, and the installation of the cables on the seabed is done by a trenching vessel. Each of these vessels has an associated chartering cost $\rho_r^V$ which is expressed per time period of 4 h. Moreover a substantial cost is also incurred for the deployment ($\rho_r^C$) and decommissioning ($\rho_r^D$) of these vessels. Also related to the (de)commissioning of the vessels are the minimal time periods for both renting and downtime. These limitations are imposed by the firm which rents out the vessel in order to guarantee that vessel occupancy rates are satisfactory. The periods during which vessels have to be rented as well as how often these vessels are (de)commissioned are determined during the simulation procedure (see Section 4.2).

## 6. Evaluating computational performance

In order to evaluate the performance of the solution techniques proposed in Section 4, two benchmarks are proposed in this

**Table 7**
Properties of the chartered vessels for offshore windfarm construction.

| Resource | Costs in 000 € | | | Renting restrictions in days | |
|---|---|---|---|---|---|
| | $\rho_r^V$ | $\rho_r^C$ | $\rho_r^D$ | Min rent period | Min down period |
| Stone dump vessel | 13.5 | 600 | 600 | 30 | 90 |
| Jack-up barge | 41.445 | 1000 | 1000 | 90 | 90 |
| Trenching vessel | 20.25 | 900 | 900 | 30 | 90 |

section. The first is an upper bound on the net present value, which is based on a best-case scenario execution of the project (Section 6.1). The second benchmark is the current deterministic scheduling approach used when scheduling similar projects (Section 6.2).

### 6.1. Upper bound for the NPV

An upper bound for the net present value earned can be calculated by making a number of assumptions regarding the manner in which the project is executed. Based on these assumptions the activity starting ($s_i$) and finishing ($f_i$) times can be obtained, which can then be used to make an estimation of the maximal possible value of the project using Eq. (1).

**Assumption 1.** All activities have their shortest possible duration. As explained in Section 3.2 activities are inherently variable, however for the upper bound it will be assumed that all activities are executed according to their most optimistic time estimate. This guarantees a minimal amount of variable costs.

**Assumption 2.** There are no delays due to weather conditions, i.e. weather conditions are always sufficiently good for the relevant activities to be executed.

**Assumption 3.** The project is executed using an as late as possible scheduling strategy without extending the critical path as defined by the shortest possible durations for all activities. By doing so all negative cashflows are shifted backwards as much as possible, while earning the positive cashflow at the end of the project as soon as possible.

Note that given the structure of the offshore project studied in this paper (see Section 5) these three assumptions guarantee continued use of the resources employed. As such, both the size and the timing of the cash outflows linked to resources is a best case scenario.

### 6.2. Deterministic scheduling

A second benchmark for the performance of the heuristics is the scheduling approach which assumes a deterministic project. This strategy uses the average durations of the activities to create a project schedule. However, noting that there may be some disruptions during the execution of the project the non-critical activities are scheduled as soon as possible (as opposed to the ALAP schedule used for the upper bound, see Section 6.1).

The performance of such a schedule can be tested using an *AG*-chromosome (Section (4.1.1). The gate times ($g_i$) for this chromosome are simply set to the starting times of this deterministic schedule. The performance as expressed by the expected net present value of the project can then be calculated using the simulation procedure described in Section 4.2.

## 7. Computational experiments

Intensive computational experiments have been carried out to investigate both the impact of specific strategies and the general performance of the different solution heuristics for the case study outlined in Section 5, as well as variations on the specific scenario with increased and decreased project size. Specifically the impact of varying the starting date of the project as well as the amount of months during which work is interrupted is investigated in Section 7.1, and more extensive experiments which compare the performance of the different algorithms for projects of various sizes are presented in Section 7.2.

### 7.1. Impact of project start and downtime

This section presents a qualitative analysis of the impact of the two main decisions by the project manager: when to start the project and how long to interrupt the project in anticipation of inclement weather. These two decisions form the basis of the dedicated heuristic algorithm which was presented in Section 4.4.3. The aim of this section is to provide insight into the impact these variables can have on the net present value of the project, based on the case study scenario with 54 wind turbines as presented in Section 5.

Naturally, the number of downtime months can only be incremented until the duration of the project extends beyond the maximal allowable deadline (which was set at five years). The experiments showed that setting the winter downtime to more than 7 months resulted in infeasible schedules, hence a total of 96 viable combinations[5] have been tested.

The net present value (*NPV*) estimate for the various strategies is summarised in Fig. 8. Looking at the leftmost panel of this figure reveals that the *NPV* displays a clear cyclical pattern when the start of the project is varied. More specifically it appears to be more beneficial to start around the month October, whereas the period surrounding May seems to be the least attractive. The difference between these two months is highly significant: starting in the right month can make a 10% difference in the net present value of the project.

The centre panel reveals a more complex pattern for the number of months the operations are halted during the winter months. The complexity of this pattern is mainly caused by the degree to which chartering costs can be decreased by not working during the winter months, taking into account the restrictions on the minimal time periods during which vessels need to be chartered (see Table 7). These results do however show that a detailed analysis of the impact of these decisions on a specific project is worthwhile since even a one month difference can have very substantial implications on the net present value of the project. Similar to the starting month decision, the choice of the number of downtime months can result in a 10% difference in the project's net present value.

A more complete picture of the solution space is presented by the surface plot in the rightmost panel of Fig. 8. This pattern clearly shows that the general patterns which have been uncovered with regard to the number of downtime months as well as the starting months remain valid when both variables are subject to variations. Nevertheless, it appears that in order to find the optimal combination it pays to perform a full factorial analysis, rather than simply basing the decision on a separate analysis of the starting month and downtime decision. Based on the separate analysis the optimal starting month would have been the 10th month (October), and the optimal number of winter downtime

[5] 12 starting months $\times$ 8 downtime month possibilities, including zero.

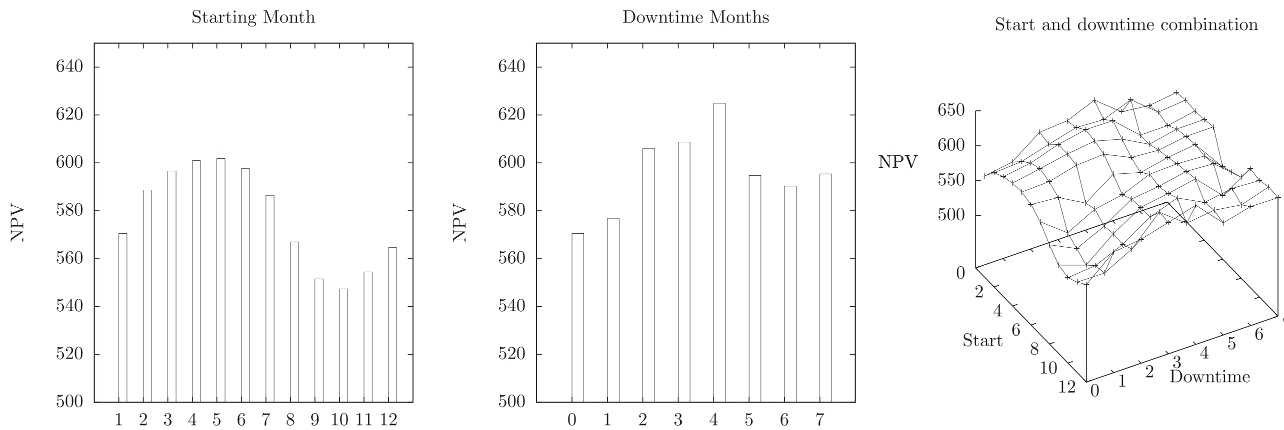


**Fig. 8.** Comparison of the impact of varying strategies on the case study project (*NPV* in mio €).

months would be four. However, when looking at the joint analysis it becomes apparent that the combination of these two decisions is sub-optimal, resulting in an estimated *NPV* of 607 million €, whereas starting in February with five months of winter downtime results in an expected *NPV* of 626 million €, an increase in *NPV* of 3.1%.

### 7.2. Comparison of algorithm performance

A second computational experiment was carried out to compare the performance of the different heuristic strategies for varying project sizes. The dataset used contained projects of a similar structure to the case study presented in Section 5 with the number of turbines to install as the determinant of project size. The number of turbines was varied between 1 and 60, each heuristic being used 20 times to get an accurate reading of the average relative performance of these heuristics. A total of 4500 h of processing time on 2.5 GHz processors was used to conduct these experiments.

Figs. 9 and 10 compare the performance of the solution approaches in absolute and relative terms respectively. The spread between the upper bound and the solution heuristics clearly increases as the problem becomes larger and systematic weather influences become more problematic.

The solution techniques which use the *SAT*-chromosome (dedicated search and *SAT*-annealing) outperform the other solution approaches, as evidenced by Fig. 10. Specifically the dedicated search significantly outperforms all other solution methods.

Interestingly, the *MAT*-annealing heuristic performs worse than the deterministic scheduling approach in the majority of the cases. This clearly shows that the solution space of this type of problem is too large to be efficiently searched at such a level of detail. This proves the usefulness of the techniques for aggregating the chromosome encoding into higher level encodings as proposed in this research. An important factor here being that these higher-level chromosome encodings still have to be translated into detailed operational guidelines (gate times $g_i$) since these are required beforehand to guarantee the availability of the installation vessels needed to complete offshore construction projects.

The relative gains from using the *SAT*-annealing technique rather than deterministic scheduling techniques is highly substantial for smaller projects. As the size of the projects increase the relative impact declines before stabilising at an increase 7–10% in expected net present value for projects with 25 turbines or more.

Both the *SAT*-annealing and the dedicated search heuristic show a dip in their relative performance for project sizes ranging from 35 to 45 turbines. The cause of this dip in performance can be seen when looking at the properties of the average solution

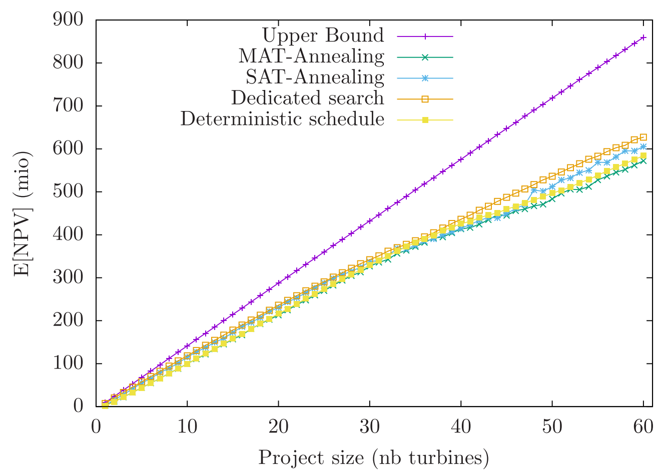


**Fig. 9.** Performance comparison in absolute terms.

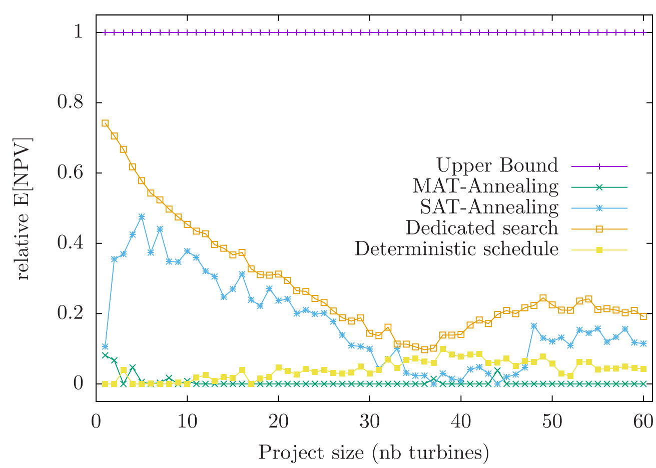


**Fig. 10.** Performance comparison in relative terms.

provided by the dedicated heuristic (Fig. 11), which clearly shows that the properties of the optimal solution change significantly within this problem size range. This implies that the increasing size of the project has caused an increased sensitivity to the seasonality of the weather patterns. For example a project spanning more than a full year will be guaranteed to be influenced by a period of bad weather, whereas a project of a couple of months can conceivably be planned around worse time periods. Moreover, as the length of the project increases the impact of discounting cashflows to calculate the net present value becomes greater.

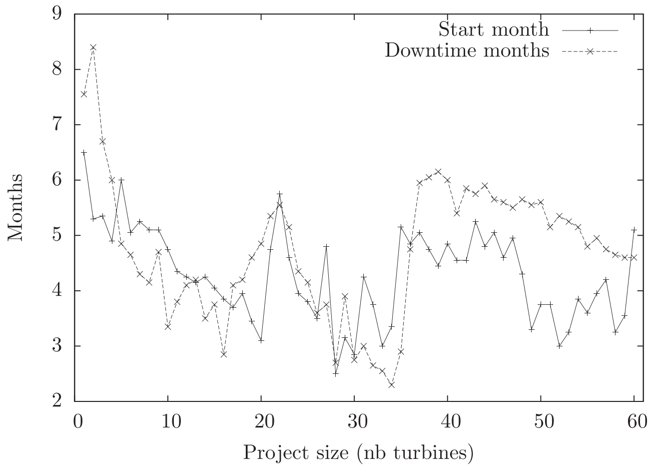


**Fig. 11.** The average properties of the optimal solution provided by the dedicated search heuristic for varying problem sizes.

## 8. Conclusions

This paper has presented a novel methodology to optimise scheduling for projects subjected to both inherent and weather-related stochasticity. This methodology consists of a novel weather model which generates both wave height and wind speed, as well as a modelling approach for the project itself.

Computational experiments indicate that the best results are obtained when using dedicated heuristics on a systematically reduced solution space. Hence, future research should focus on efficient ways of identifying the most interesting areas of the solution space for this type of problem. Moreover, future research should extend the analysis to include other case studies, as well as creating a dataset containing benchmark problems on which future heuristics can be compared objectively. The discussion of other case studies as well as climates could also be a valuable extension of this research. For projects where the lead times for resources are short or non-existent a dynamic scheduling approach may also be an interesting extension of this research.

## Acknowledgements

We acknowledge the support provided by the Bijzonder Onderzoeksfonds (BOF) for the project with Contract number BOF12-GOA021. This work was carried out using the STEVIN Supercomputer Infrastructure at Ghent University, funded by Ghent University, the Flemish Supercomputer Center (VSC), the Hercules Foundation and the Flemish Government department EWI.

## Appendix A. Weather model persistence plots

As was noted in the literature review (see Section 2), the persistence of weather conditions is an important indicator of the quality of a weather model. Moreover, this aspect of the weather model is also of great importance to accurately estimate the influence of weather conditions on activity durations.

Figs. A1 and A2 show how the persistence of wind- and wavestates of the weather generation model compare to the actual persistence observed from the empiric data gathered from the offshore measuring stations. Each of these figures shows a separate

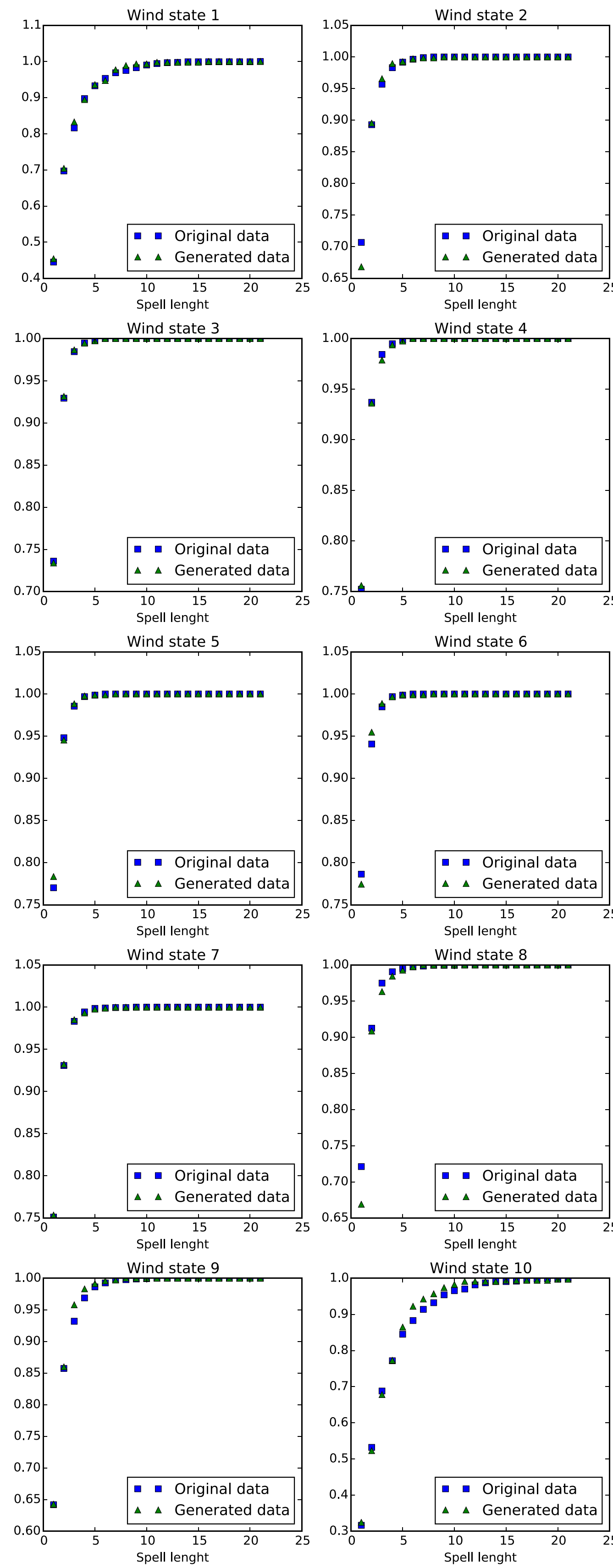


**Fig. A1.** Wind state persistence.

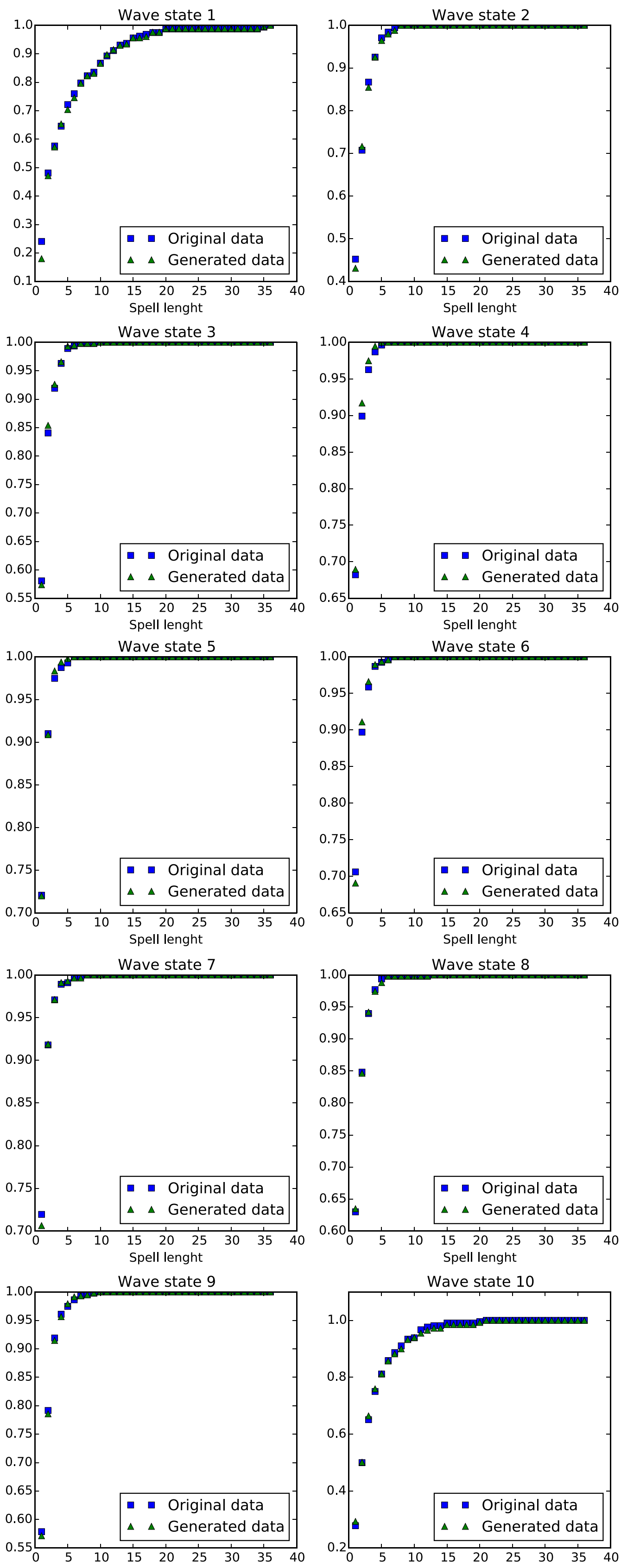


**Fig. A2.** Wave state persistence.

chart for each of the ten possible states for wind and waves respectively. These charts show the cumulative probability of remaining in a certain wind- or wavestate for a certain number of periods. From these charts it can be concluded that the simulation model is capable of modelling the persistence of both wind- and wave conditions with great accuracy.

## References

[1] Achuthan N, Hardjawidjaja A. Project scheduling under time dependent costs? A branch and bound algorithm Annals of Operations Research 2001:55–74.

[2] Ailliot P, Monbet V, Prevosto M. An autoregressive model with time-varying coefficients for wind fields. Environmetrics 2006;17(September (2005)): 107–117.

[3] Arena F, Puca S. The reconstruction of significant wave height time series by using a neural network approach. Journal of Offshore Mechanics and Arctic Engineering 2004;126(3):213.

[4] Ballestín F. A genetic algorithm for the resource renting problem with minimum and maximum time lags. In *Lecture notes in computer science*, vol. 4446; 2007. p. 25–35.

[5] Ballestín F. Different codifications and metaheuristic algorithms for the resource renting problem with minimum and maximum time lags. In: Cotta C, van Hemert J, editors. Recent advances in evolutionary computation for combinatorial intelligence of studies in computational intelligence, vol. 153. Berlin, Heidelberg: Springer-Verlag; 2008. p. 187–202 [Chapter 12].

[6] Bendavid I, Golany B. Setting gates for activities in the stochastic project scheduling problem through the cross entropy methodology. Annals of Operations Research 2011;189(1):25–42.

[7] Bowers J, Mould G. Weather risk in offshore projects. Journal of the Operational Research Society 1994;45(4):409–418.

[8] Brokish K, Kirtley J. 2009. Pitfalls of modeling wind power using Markov chains. In: Power systems conference and exposition, 2009. PSCE'09, Seattle. IEEE/PES.

[9] Cheng T, Ding Q, Lin B. A concise survey of scheduling with time-dependent processing times. European Journal of Operational Research 2004;152(1): 1–13.

[10] Cunha C, Guedes Soares C. On the choice of data transformation for modelling time series of significant wave height. Ocean Engineering 1999;26(6):489–506.

[11] Etgar R, Shtub A, LeBlanc L. Scheduling projects to maximize net present value—the case of time-dependent, contingent cash flows. European Journal of Operational Research 1996;96:90–96.

[12] Ettoumi F, Sauvageot H, Adane aE-H. Statistical bivariate modelling of wind using first-order Markov chain and Weibull distribution. Renewable Energy 2003;28(11):1787–802.

[13] European Wind Energy Association. Wind in our sails—the coming of Europe's offshore wind energy industry. Technical report; 2011.

[14] Fahmy A, Hassan T, Bassioni H. What is dynamic scheduling? PM World Journal 2014;III(V):1–9.

[15] Goldratt E. Critical chain. Great Barrington, MA: Press; 1997.

[16] Graham C. The parameterisation and prediction of wave height and wind speed persistence statistics for oil industry operational planning purposes. Coastal Engineering 1982;6(4):303–329.

[17] Hartmann S, Briskorn D. A survey of variants and extensions of the resource-constrained project scheduling problem. European Journal of Operational Research 2010;207:1–15.

[18] Herroelen W, Leus R. The construction of stable project baseline schedules. European Journal of Operational Research 2004;156(3):550–565.

[19] Herroelen W, Leus R. Project scheduling under uncertainty: survey and research potentials. European Journal of Operational Research 2005;165: 289–306.

[20] Herroelen W, Van Dommelen P, Demeulemeester E. Project network models with discounted cash flows a guided tour through recent developments. European Journal of Operational Research 1997;100:97–121.

[21] Huang W, Ding L. Project-scheduling problem with random time-dependent activity duration times. IEEE Transactions on Engineering Management 2011;58(2):377–387.

[22] Ke H, Liu B. Project scheduling problem with stochastic activity duration times. Applied Mathematics and Computation 2005;168(1):342–353.

[23] Kerkhove L-P, Vanhoucke M, Maenhout B, Vandenheede L. On the resource renting problem with overtime. Working paper at Ghent University, currently under journal submission (Check ⟨www.projectmanagement.ugent.be⟩ *for an update*); 2015.

[24] Makarynskyy O, Pires-Silva aa, Makarynska D, Ventura-Soares C. Artificial neural networks in wave predictions at the west coast of Portugal. Computers& Geosciences 2005;31(4):415–24.

[25] Mandal S, Prabaharan N. Ocean wave prediction using numerical and neural network models. The Open Ocean Engineering Journal 2010;3(1):12–17.

[26] Mehta S, Uzsoy R. Predictable scheduling of a job shop subject to breakdowns. IEEE Transactions on Robotics and Automation 1998;14(3):365–378.

[27] Möhring R, Schulz A. Solving project scheduling problems by minimum cut computations. Management Science 2003;49(3):330–350.

[28] Möhring R, Schulz A, Stork F, Uetz M. On project scheduling with irregular starting time costs. Operations Research Letters 2001;28(664):149–154.

[29] Monbet V, Ailliot P, Prevosto M. Survey of stochastic models for wind and sea state time series. Probabilistic Engineering Mechanics 2007;22(2):113–126.

[30] More A, Deo M. Forecasting wind with neural networks. Marine Structures 2003;16(1):35–49.
[31] Norlund EK, Gribkovskaia I, Laporte G. Supply vessel planning under cost, environment and robustness considerations. Omega 2015;57:271–281.
[32] Nübel H. The resource renting problem subject to temporal constraints. OR Spektrum 2001;23:359–381.
[33] Qian F, Gribkovskaia I, Laporte G, Halskau sr Ø. Passenger and pilot risk minimization in offshore helicopter transportation. Omega 2012;40(5):584–593.
[34] Qian F, Strusevich V, Gribkovskaia I, Halskau Ø. Minimization of passenger takeoff and landing risk in offshore helicopter transportation: models, approaches and analysis. Omega 2015;51:93–106.
[35] Regnier E. Doing something about the weather. Omega 2008;36(1):22–32.
[36] Rothkopf M, McCarron J, Fromovitz S. A weather model for simulating offshore construction alternatives. Management Science 1974;20(10):1345–1349.
[37] Sahin AD, Sen Z. First-order Markov chain approach to wind speed modelling. Journal of Wind Engineering and Industrial Aerodynamics 2001;89(3–4): 263–269.
[38] Schönefeld P. Private communication; 2014.
[39] Scotto M, Guedes Soares C. Modelling the long-term time series of significant wave height with non-linear threshold models. Coastal Engineering 2000;40 (4):313–327.
[40] Shamshad a, Bawadi M, Wanhussin W, Majid T, Sanusi S. First and second order Markov chain models for synthetic generation of wind speed time series. Energy 2005;30(5):693–708.
[41] Sobel MJ, Szmerekovsky JG, Tilson V. Scheduling projects with stochastic activity duration to maximize expected net present value. European Journal of Operational Research 2009;198(3):697–705.
[42] Stefanakos C, Athanassoulis G. Multiscale time series modelling of significant wave height. In: Proceedings of the twelfth (2002) international offshore and polar engineering conference, vol. 3; 2002. pp. 66–73.
[43] Stefanakos CN. Nonstationary stochastic modelling of multivariate long-term wind and wave data. In: Proceedings of the 24th international conference on offshore mechanics and Arctic engineering (OMAE 2005), number Omae; 2005.
[44] Stork F. Branch-and-bound algorithms for stochastic Resource-constrained project scheduling. Technical report, (702/2000); 2000.
[45] Talbi E-G. Metaheuristics: from design to implementation; 2009.
[46] Trietsch D. Optimal feeding buffers for projects or batch supply chains by an exact generalization of the newsvendor result. International Journal of Production Research 2006;44(4):627–637.
[47] Vandenheede L, Vanhoucke M, Maenhout B. A scatter search for the extended resource renting problem; 2014.
[48] Vanhoucke M. Work continuity constraints in project scheduling. Journal of Construction Engineering and Management 2006;132:14–25.
[49] Vanhoucke M. Project management with dynamic scheduling: baseline scheduling, risk analysis and project control, vol. XVIII. New York: Springer; 2012.
[50] Vanhoucke M, Demeulemeester E, Herroelen W. Maximizing the net present value of a project with linear time-dependent cash flows. International Journal of Production Research 2001;39:3159–3181.
[51] Wilks DS, Wilby RL. The weather generation game: a review of stochastic weather models. Progress in Physical Geography 1999;23(3):329–357.