1. Introduction

An important factor related to the marine environment is wave conditions [1]. Information regarding wave height is very important for designing and planning activities in coastal, e.g. port, as well as offshore [2]. Large sea waves may pose a significant risk to ship movement, coastal structure, and operations at the port. The development of activities in ocean waters in nearshore as well as offshore, need the information of wave condition [3].

Wave height prediction is needed to design and operate many projects in marine waters especially for hazard mitigation. From navigation and civil protection, reliable measurement of events in ocean waters is very important [3]. Wave data is difficult to obtain and computationally expensive compared to wind measurements, therefore wave data can be obtained indirectly from wind measurements [4]. Unfortunately, both the wind and wave measurement data are very rare and are very expensive to obtain via measurement.

Researchers have been developing various methods to obtain accurate wave predictions. The traditional way approach is via semi-empirical models that are relatively fast and low-cost computation, such as the SMB method [5], Shore Protection Manual [6], and Coastal Engineering Manual [7]. Here, the wave prediction is based on fetch length, wind speed, and duration, in which the wave directions are assumed to be the same as the wind direction. Another approaches, the wave prediction are obtained via numerical wave simulation by using the 3rd spectral wave models such as SWAN [8] and WAM [9]. But, due to the high computational cost, the numerical approach is relatively difficult to perform.

Another approach for wave height prediction is by using soft computing. In 2011, Malekmohamadi et al. [4] performs research that discussed several soft computing-based models. Here, wind and wave height data are obtained from NDBC or the National Buoy Center, in Lake Superior, USA. Three methods were compared, i.e. the Artificial Neural Network (ANN), the Support Vector Machine (SVM), and the Adaptive Neuro-Fuzzy Inference System of ANFIS [4]. Recently, in 2018, James et al. [10] uses some machine learning approaches such as Multi-Layer Perceptron (MLP) and SVM to replace traditional numerical simulation for predicting wave condition, especially wave height and wave periods. Here, numerical simulation with SWAN and Wave Watch III (WW3) are used to build training data for predicting wave height and wave period.

This paper aims to use a soft-computing approach by using Support Vector Regression (SVR), especially for reducing time computing. Since the limitation of wind and wave observation, in this study, we use the wind data based on the ECMWF ERA-5 [9], whereas the wave data for training data is obtained by SWAN model [8] via numerical simulation. The training data obtained are then used to train the Support Vector Regression (SVR) model. This research will discuss the steps to build training data, implement Support Vector Regression (SVR) using several kernels, and measure the performance of the SVM model for wave height prediction. As a study case, we choose an area with rather complex geometry and bathymetry, i.e. archipelago country consisting of small islands, i.e. Jakarta Bay, in Indonesia.