

# Phase-resolved Wave Prediction from a SWIFT Buoy Array

Michael Schwendeman

30 April, 2017

## 1 Introduction

This project is motivated by a long-standing problem in wave-energy generation. It is known that wave energy converters (WECs) can produce significant gains in efficiency by using advanced control techniques. However, such techniques often require advanced knowledge of the incoming waves (see, for example, Hals et al., 2011). In the ideal case, the system is given an exact trace of the future wave excitation force. As this is a difficult problem, most WECs are tuned instead based only on the average 30-minute bulk wave parameters.

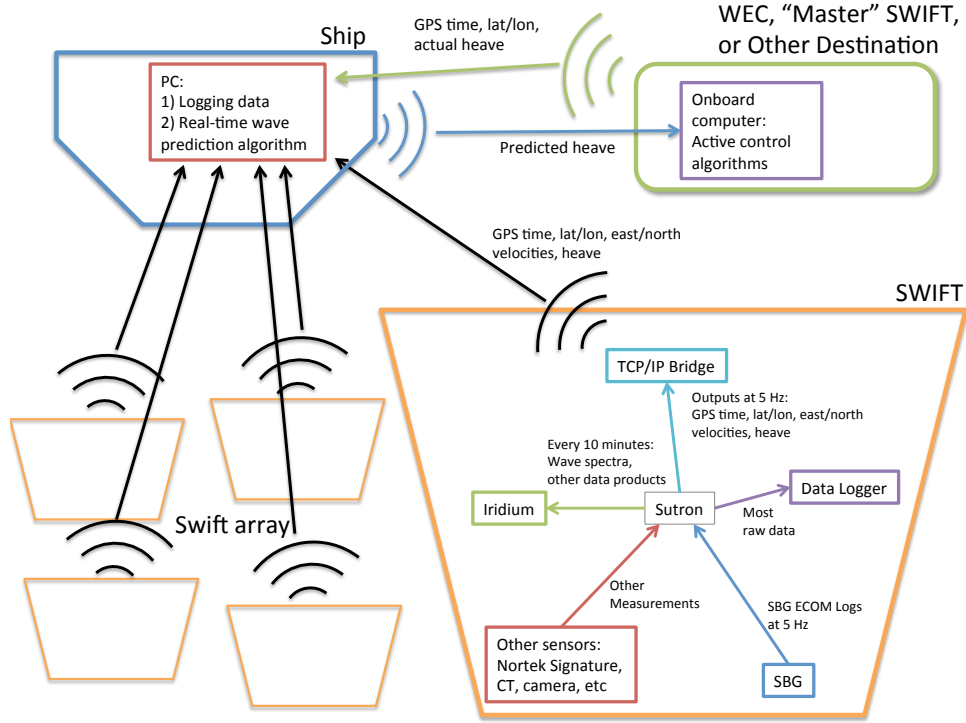
The goal of this project is to provide a useful prediction of the incident waves on a WEC, using an array of networked buoys. By useful we mean that when coupled with advanced control techniques, the wave prediction improves upon the WEC performance over just using average wave parameters. A related question, which may have widespread scientific applicability, is whether we can construct a realistic approximation of the sea surface from the sparse buoy array.

This report documents the progress made so far on these questions. It is broken into three sections. Section 2 outlines the integration of the SBG Ellipse IMU into the newest version (“v4”) SWIFT buoys, and describes the Python program for receiving data in real-time using TCP-IP sockets and an RF ethernet bridge. Section 3 details the development of a phase-resolved wave prediction algorithm using linear sea surface simulations of directional spectra. Section 4 shows a preliminary evaluation of the phase-resolved algorithm using real data obtained from a research cruise off of Southern California. As this is very much a work in progress, each section includes several recommendations and ideas for future work, which are together summarized at the end of the report.

## 2 SBG and Ethernet Bridge Integration

### 2.1 System Overview

A schematic of the buoy array approach to wave prediction is shown in Figure 1. One of the major design considerations of the new SWIFT (surface wave instrument float with



**Figure 1: Schematic showing the main components of the SWIFT buoy array and wave prediction architecture.**

tracking) v4 drifters was their use in the wave prediction problem. As in the version 3 SWIFTs, the SWIFT v4 is built around a Sutron Xpert microprocessor and data logger. The Xpert has a 32-bit processor, supports SD memory, and has multiple serial and ethernet ports. The Xpert logs the raw measurements from the SWIFT sensor payload, and performs onboard processing of certain data products. These products are then transmitted over Iridium satellite communication to a server on land once per hour.

The form factor of v4 buoys is significantly different than that of the previous SWIFTs, to accommodate the down-looking Nortek Signature 5-beam ADCP. The v4 buoys no longer have the tall spar-type shape of the previous SWIFTs, which makes them slightly more susceptible to tipping motions. Ballast was added to the buoy hulls below the water line to limit this effect, and drogues have been experimented with as well.

For the purposes of this project, the most important change in the SWIFT architecture is the transition from a MicroStrain inertial package to one made by SBG Systems. The SBG Ellipse-N is a miniature, high-performance MEMS based inertial system with an industrial GNSS receiver. The Ellipse-N inertial motion unit (IMU) features three-axis

gyroscopes, accelerometers, and magnetometers. It also runs a real-time Extended Kalman Filter (EKF) for data fusion of its IMU and GPS measurements. In combining the IMU and GPS data streams, the EKF uses the advantages of each measurement to produce a better position and orientation solution than from either measurement alone.

The most critical measurement of the SWIFTs for real-time wave prediction is the buoy vertical position, or heave. This is a difficult measurement because it requires two integrations of the inertial accelerometer signal, which leads to drift in the position. At the same time, the accuracy of the GPS measurements is much worse in the vertical direction than horizontal. To solve this problem, SBG have developed an advanced high pass filter specifically intended for estimating heave in real-time for marine purposes. The filter is designed to limit phase and gain errors in wave motions with periods shorter than 15 seconds.

## **2.2 SBG Ellipse Configuration and Magnetometer Calibration**

The SBG Ellipse-N is configured and calibrated using the sbgCenter application, which requires the device to be plugged in through serial or usb connectors. sbgCenter allows you to set the output logs, connection settings, device orientation, lever arms, and aiding sensors (such as GPS). The device saves its most recent configuration, so this step does not need to be repeated for each deployment. Configuration settings are also saved and can be loaded from a file. See the dropbox SBG Systems folder in the SWIFT\_v4.x directory for the most up-to-date configuration.

One additional step is needed to ensure accurate measurements from the Ellipse, which is calibration of the internal magnetometer. This is necessary because although the Ellipse comes calibrated from the factory, placing the Ellipse in the SWIFT puts it near components which may distort the magnetic field. In this way the uncalibrated Ellipse heading measurement may be off by tens of degrees, which would seriously effect the accuracy of many of the necessary data products. The calibration procedure is described in the Ellipse Hard & Soft Iron Calibration Manual. Unfortunately, the calibration uses the sbgCenter application, which requires the device to be plugged in to the computer. Thus, the SWIFT cannot be fully sealed, which leaves open the possibility for small errors. Caveats aside, the calibration procedure appears to bring the average heading error down from roughly 5% or more to 1-2%.

## **2.3 Receiving SBG Data Over Ethernet Bridge**

The SBG Ellipse-N data is stored and transmitted as messages using the "sbgECom" binary protocol, as described in the Firmware Reference Manual. The binary messages are logged to a file by the Xpert processor, and also transmitted asynchronously over a TCP/IP socket bridge. The TCP/IP signal is then wirelessly transmitted over RF radio (900 MHz) via the Digi XPress Ethernet Bridge. The XPress Ethernet Bridge has a stated

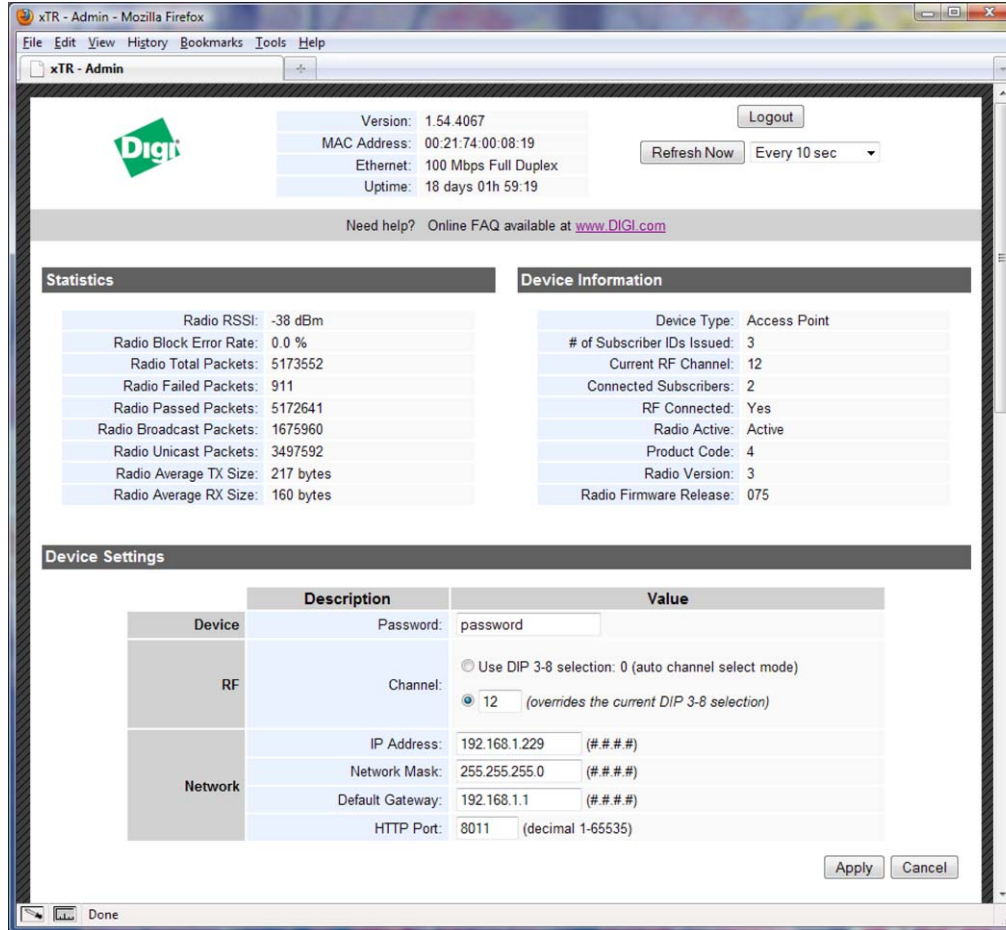


Figure 2: Example of the Digi Xpress Ethernet Bridge browser interface.

range of 2-15 miles depending on the antenna used, and a data rate of 1.5 Mbps. In Figure 1, the data from the SWIFT array and WEC (or other target point) are sent to a shipboard PC, where the prediction algorithm is run and returned to the WEC for active controls. Alternatively, it may be preferable to have the algorithm running in an embedded system onboard the WEC.

To stream data over the ethernet bridge to the PC, first follow the directions in the file Xpert\_EthernetBridge.setup.txt. This sets the IP addresses, Mac addresses, ports, etc. for both sides of the bridge — namely the shipboard PC and Xpert microprocessor on the SWIFT. To check that the Ethernet Bridge is setup properly, simply open a browser and navigate to the device IP address, where a screen will show statistics of the connection and device information, as shown in Figure 2.

The software for receiving data from the SWIFTs and decoding the SBG binary messages is written in Python and contained in two files, which are outlined below. The file `readAndDecodeFromEthernetBridge.py` is a script which, when run from the PC, opens a TCP/IP socket, listens for incoming connections, and accepts data as available. In this way, the PC acts as a server, in that it does not query for data, but accepts the data asynchronously. The TCP/IP functionality is accessed through the Python socket module, which is available on all modern Unix, Windows, and MacOS platforms.

As the data streams in, a series of if-statements check for the sync bytes that indicate the beginning of an sbgECom message, as detailed in the Ellipse Firmware Reference Manual. When such a message is detected, the script calls a function `parseSbgMessage` which further decodes the binary data. The parsing functions are accessed through the `sbgMessageParse` module contained in the file `sbgMessageParse.py`. The function `parseSbgMessage` was written such that it can read binary data from a socket connection, as indicated above, or from a previously saved raw binary file. In addition, it can print decoded data in ascii format to a file or to stdout. The default behavior is to read the data from a socket connection and not print the output.

A key piece of the parsing is the dictionary variable `sbgMessages`. The `sbgMessages` dictionary contains all the information from the Firmware Reference Manual on the message ID, message name, message length (in bytes), field names, and field types (e.g. 8-bit unsigned integer, double float, etc.). For example, take the first entry:

```
sbgMessages = {b'\x01': {'name': 'Status', 'intLength': 22, '
    binLength': b'\x16\x00', 'unpackString': '<LHHLLH', 'fields': ( '
    time_stamp', 'general_status', 'reserved1', 'com_status', '
    aiding_status', 'reserved2', 'reserved3' ) ,...
```

This indicates that messages with `msgID=1` (b'\x01' in Python binary syntax) has name 'Status,' length of 22 bytes (b'\x16\x00' in binary), and fields 'time\_stamp' (unsigned 32 bit integer, or "L"), 'general\_status' (unsigned 16 bit integer, or "H"), etc. All of this information is taken from the Ellipse Firmware Reference Manual. This format makes it easy to add entries to the `sbgMessages` dictionary if new SBG messages are required at a later time.

The function `parseSbgData` is called by the wrapper function `parseSbgMessage`, and is responsible for actually translating the binary messages into a usable Python data structure. The code is quite compact:

```
### Function for parsing binary SBG messages into dictionaries
def parseSbgData(msgID, binData):
    parsedData = struct.unpack(sbgMessages[msgID][ 'unpackString '
        ], binData)
    dataStruct = OrderedDict( zip(sbgMessages[msgID][ 'fields' ],
        parsedData))
```

```
return dataStruct
```

It uses the `unpack` function of the `struct` library to translate the message bytes into a tuple of integers and floats, as defined in the `unpackString` of the `sbgMessages` dictionary. Then, it puts those values into an ordered dictionary with the field names for that message type, also defined in the `sbgMessages` dictionary. This `dataStruct` ordered dictionary is the final result of the parsing and is returned and/or printed. The Python ordered dictionary works just like a regular dictionary except that when printed, the fields are always in order (i.e. with `time_stamp` always printing first).

Finally, it should be noted that streaming and decoding of the SBG data could alternatively have been performed using C++. Specifically, C comes with its own sockets library, and SBG provides an API, written in C++, which can decode the binary messages. Python was used as it provided a more user-friendly approach, more online support, and easier debugging. Additionally, a similar decoder was written in Matlab (function `sbgBinaryToMatlab.m` and script `sbgBinaryToMatlab_Batch.m`) for use with logged binary files. For example, these Matlab functions are used to decode the binary data for testing the prediction algorithm in Section 4.

## 2.4 Future Work

Some additional work is needed to operationalize the above procedure. Notably, the Ethernet Bridge has not been tested in long range outside, and has not been installed permanently on the SWIFT hulls. The system has been tested in the laboratory however, and works well to capture and decode SBG data from a single SWIFT over the Ethernet Bridge.

The `readAndDecodeFromEthernetBridge.py` program needs to be written to accommodate multiple SWIFTs. This may be slightly tricky, as the current script is quite linear. It seems that rather than wait for one connection, it could be modified to wait for four (or as many SWIFTs are in use). Meanwhile, in the `while True` loop, it could receive bytes from each connection and parse them one after another. However, this flow has not yet been tested. Alternatively, a more complex program would make use of the asynchronous processing modules available in Python, such as `asyncio`.

A helpful feature of the system is that data is saved in a buffer until the PC is ready to receive the data, so bytes are not lost. Along the same lines, one feature that has not been implemented is a cyclic redundancy check (CRC), which can be used to look for dropped bytes in the data. The outline of such a function is given in the SBG Firmware Reference Manual.

With the above modifications, the system will be able to run on a shipboard PC or embedded system onboard a WEC, allowing it to receive, parse, and save timestamped location and heave data from multiple SWIFTs over the RF Ethernet Bridge. At which point, the next step will be to implement an algorithm which takes in that data, and

outputs a prediction (in real time) of the incoming waves to the WEC. The next sections outline the progress being made in the development of such an algorithm.

### 3 Linear Simulations and Least Squares Prediction Algorithm

#### 3.1 Least Squares Prediction Algorithm

Our goal is to develop an algorithm which, given a sparse array of wave measurements, can predict (in some way) the waves that will pass a different point at a later time. Connell et al. (2015) describe one such algorithm, which is based on a least squares fit of the measurements to different wave components, which are then propagated forward in time by assuming linear dispersion. Given enough observations and minimal nonlinearity, this method yields a highly accurate time-domain, phase-resolved forecast of the approaching waves. The Connell et al. (2015) paper is written around the use of a coherent Doppler radar for making the wave observations. The advantage of radar over a sparse buoy is in the much larger number of observations, with much better spatial coverage. One downside is that the radar can only estimate the radial wave orbital velocity, rather than surface elevation, which slightly complicates the analysis.

The algorithm is based around the representation of the sea surface as a sum of individual wave components, as in

$$\eta(\mathbf{x}, t) = \Re \left( \sum_{n=1}^N A_n \exp(i(\mathbf{x} \cdot \mathbf{k}_n - \omega_n t)) \right) \quad (1)$$

where  $A_n$ ,  $\mathbf{k}_n$ , and  $\omega_n$  are the amplitude, wavenumber, and frequency of each component. Assuming linear dispersion, then gives

$$\omega = \sqrt{gk} \quad (2)$$

where  $g$  is gravitational acceleration and

$$k = \sqrt{k_x^2 + k_y^2}. \quad (3)$$

This equation can be expressed in matrix form as

$$\eta = \mathbf{P}\mathbf{A} \quad (4)$$

where  $\mathbf{P}$  is an  $N \times M$  matrix with elements

$$P_{n,m} = \exp(i(\mathbf{x}_m \cdot \mathbf{k}_n - \omega_n t_m)), \quad (5)$$

called the propagator matrix.  $\eta$  is an  $M \times 1$  vector of observations, and  $\mathbf{A}$  is an  $N \times 1$  vector of unknown wave amplitudes. If  $M > N$ , this is an overdetermined system of equations, such that  $\mathbf{A}$  can be fit using the linear least squares method. With the component amplitudes determined in this way, the surface elevation at a desired time and place can be projected using equation 1

The least squares prediction algorithm is implemented in Matlab in the function `leastSquaresWavePropagation.m`. Below is a slightly simplified version of the code:

```
function [z2,zc] = leastSquaresWavePropagation_2D(z1,t1,x1,y1,t2,
    x2,y2,k,theta,reg_factor)

kx = k*cos(theta');
ky = k*sin(theta');
omega = sqrt(9.8*k)*ones(size(theta'));
P1 = [cos(x1*kx'+y1*ky'-t1*omega'),sin(x1*kx'+y1*ky'-t1*omega')]';
A = ridge(z1,P1,reg_factor,0);
P2 = [cos(x2*kx'+y2*ky'-t2*omega'),sin(x2*kx'+y2*ky'-t2*omega')]';
z2 = P2*A(2:end);
end
```

Note that the code above shows two slight differences from how the algorithm was first described. First, is that the wave components are expressed as sines and cosines, rather than the real part of a complex exponential. In theory, these representations are exactly the same. However for a reason that is not clear, when complex exponentials are used, the fitted result is off by a factor of two. This can be readily tested with simple examples using a surface constructed of only a few known components.

The second difference is more conceptually important. It was found often that using a pure least squares solution led to overfitting, thus it was replaced with a regularized least squares method, also called ridge regression. Overfitting produces large values for the wave component amplitudes, which largely cancel out to reproduce the surface elevation at the observed points, but leads to unphysical surface elevations everywhere else. Regularization penalizes large values of fitted coefficients, leading to more realistic results. However, it also introduces a new parameter, `reg_factor`, which controls the degree of penalization.

### 3.2 Simulations of Irregular Waves from Directional Spectra

This problem is aided greatly by the use of synthetic data to test and refine the forecast algorithm. Thus, the Wave Analysis for Fatigue and Oceanography (WAFO) toolbox is used to produce realistic simulations of irregular ocean waves in three dimensions (Brodtkorb et al., 2000). Specifically, the function `seasim.m` is run, which uses the discrete spectral approximation of Dietrich and Newsam (1997) to efficiently simulate 3D surfaces from a given directional wave spectrum.



The wave simulations are carried out in the script `simulateWaveSurfaceFromDirectionalSpectrum.m`. Two inputs are needed from the user. The first is a source for the directional energy spectrum. Three examples are given: SWIFT and Datawell Waverider spectra from Station Papa, and CDIP data from San Nicolas Island. Regardless of the source, the spectrum must be put in WAFO's specified wave spectra structure format, as described in the file `spec.pdf` in the WAFO toolbox directory. For example, the wave directions must be in radians and range from  $-\pi$  to  $\pi$ . Optionally, the script can rotate the directional spectrum such that the peak wave direction (direction energy is propagating from, not to) is aligned with the  $+x$  coordinate. Next, the user must define the simulation grid, including spatial and temporal range ( $L$  and  $T$ , in meters and seconds respectively) and resolution ( $dx$  and  $dt$ ). Note that the simulations run slowly and produce large arrays of data when given overly large domains.

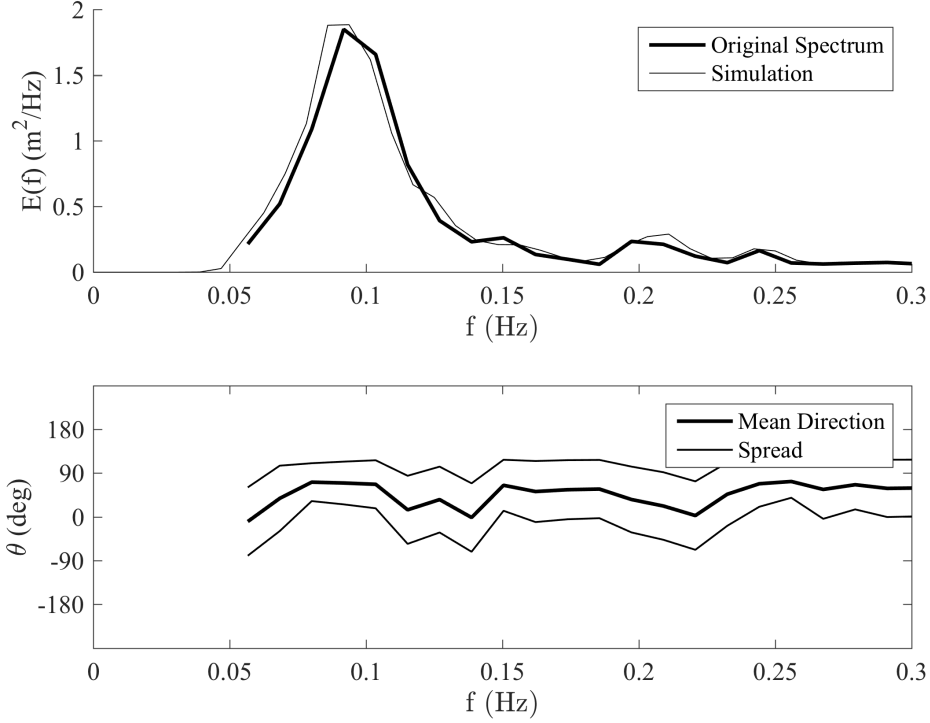
The `seasim` function takes as input the directional spectrum (in WAFO format), and grid dimensions. It also takes an `fftdim` parameter, which defines whether the simulation runs a two-dimensional inverse FFT at each timestep (`fftdim=2`) or a one-dimensional inverse FFT at each spatial grid cell (`fftdim=1`). When spectra are made from the simulation data, they appear to match the original (input) spectra better when using the one-dimensional inverse FFT method. `seasim` outputs the simulation results as a structure, `Y`, which has fields `x`, `y`, `t`, and `Z`. `simulateWaveSurfaceFromDirectionalSpectrum.m` finishes by saving the structure `Y`, and plotting the input directional spectrum and a check of the original frequency spectra with one calculated from the simulated data. These usually match well, particularly for large simulations, as shown in Figure 3.

### 3.3 Algorithm Performance for Simulated Data

- Two functions, one to calculate wave prediction at a point, one to make a gridded surface reconstruction.
- Input parameters: buoy number and array shape, target location or grid size and resolution, wave components and regularization parameter, time windows and delay
- Several examples. Spectra source. Buoy shape. Buoy number. Regularization and buoy components. Time window.
- Runtime results

### 3.4 Future Work

The simulations show that the least squares forecast algorithm can work well with only a limited number of buoys under ideal conditions, namely, long-crested, energetic swells with little wind sea. Moreover, they show that there is significant sensitivity based on the spatial location of the buoy array, relative to the wave conditions and the parameters of the least squares fit. Clearly, there are many questions that the synthetic data cannot



**Figure 3: Example comparison of frequency spectra from `simulateWaveSurfaceFromDirectionalSpectrum.m`. Data from a SWIFT at Station Papa,  $L = 500$  m,  $T = 300$  s,  $dx = 4$  m,  $dt = 1$  s.**

address, such as the complicating factor of nonlinearity. However, to better inform future field experiments, a systematic analysis of the relationship of the different parameters to the forecast error would be useful. For example, how does the error scale with the ratio of the buoy array separation to the peak wavelength? How many buoys are needed, and does it change with different wave conditions? Perhaps most critical is understanding the time window for making a valid forecast, and how that depends on the other parameters. There is some work being done on this subject in the laboratory of Dick Yue at MIT, particularly his graduate student Yusheng Qi, although it is as yet unpublished.

These simulations also provide a convenient laboratory for experimenting with the forecast algorithm. One promising idea for improving the algorithm is to add to the least squares cost function a soft constraint based on the directional wave spectra. Since on average the wave component amplitudes will match the measured spectra, the best guess of the amplitudes for any one window of data would also match the spectra. This would amount to providing an individual `reg_factor` for each component of the fit, although it is not immediately clear how the spectral values would translate into the cost function values

(for example, are they proportional?). Some care would be needed in implementing this strategy.

## 4 Evaluation from SWIFT data

- Describe LC-DRI experiment.
- Example SWIFT wave prediction.
- Envelope calculations.
- Examine surface reconstruction.
- Future work: rolling algorithm with incrementation, use velocity data

## 5 Summary of Future Work

- So far IO software integration and algorithm development progressed separately.
- Hardware and IO integration could be ready for testing soon. Runs successfully on laboratory PC. Has not been tested outside, running on embedded system, with buoys at long range. Currently only written for one buoy.
- Algorithm development could use work. Rolling estimator constrained by directional spectra would help against over-fitting to high frequencies and jumps between windows. Use of u,v velocities might be more accurate in field data.
- Systematic error analysis would be useful for informing future experiments. How many buoys are needed and how close together must they be? How to relate the distance between array and target to the time window and delay?

## References

- Brodtkorb, P. A., P. Johannesson, G. Lindgren, I. Rychlik, J. Rydén, and E. Sjö (2000). Wafo - a matlab toolbox for the analysis of random waves and loads. In *Proc. 10'th Int. Offshore and Polar Eng. Conf., ISOPE, Seattle, USA*, Volume 3, pp. 343–350.
- Connell, B. S. H., J. P. Rudzinsky, C. S. Brundick, W. M. Milewski, J. G. Kusters, and G. Farquharson (2015). Development of an environmental and ship motion forecasting system. *Proceedings of the ASME 2015 34th International Conference on Ocean, Offshore and Arctic Engineering*.
- Dietrich, C. R. and G. N. Newsam (1997). Fast and exact simulation of stationary gaussian processes through circulant embedding of the covariance matrix. *SIAM J. Sci. Comput.* 18(4), 1088–1107.

Hals, J., J. Falnes, and T. Moan (2011). Constrained optimal control of a heaving buoy wave-energy converter. *J. Offshore Mech. and Arctic Eng.* 133.