Comparing the Variability of Wind Speed and Wave Height Data

Tom Denniss

Energetech Australia Pty Ltd, Sydney, Australia

Abstract

A comparison of wind speed and wave height distributions is performed using a simplified probability distribution function similar to the Weibull distribution. In general, wind and wave data can be accurately fitted to this theoretical distribution. The distribution was specifically chosen because the higher moments of the distribution are easily manipulated mathematically, thus resulting in simple analytical representations for the associated power distributions. While the relative variability (defined as the ratio of the standard deviation of the distribution to its mean) of wind speeds and wave heights is generally comparable, this is not the case for their respective power distributions, which exhibit a greater variability for wind power than for wave power. This is due largely to the fact that, given an equivalent variability of the underlying distributions, wind power is a function of the cube of wind speed, while wave power is primarily a function of the square of wave height. In this sense, a wind power distribution is more "stretched out" than a wave power distribution. This outcome results in a much higher "capacity factor" for wave energy, compared to that of wind energy.

1. Introduction

Renewable energy sources are receiving much more attention now than even a few years ago. Wind and wave energy are two such sources and technologies based on these forms of energy have been developed seriously in recent times. Wind energy is generally considered to be about twenty years ahead of wave energy in its stage of development.

Most forms of renewable energy implicitly entail variability. This variability has been a problem for developers, as energy retailers require a constant or, at least, a very predictable source of energy in order to minimise the need for reserve capacity on their electricity grid. Wind energy has come in for much criticism from some quarters for its supposed high level of variability. Wave energy will not be immune to this same criticism, although its higher level of predictability will reduce the need for spinning reserve capacity.

This paper attempts to construct a framework which demonstrates that wave energy is, in fact, substantially less variable than wind energy. This is not to say that wind speeds themselves are necessarily more variable than wave heights and periods. However, the functional relationship between wind power and the cube of the wind speed contrasts somewhat with the corresponding relationship between wave power and the square of wave height. Wave power also relies, to a much lesser degree, on wave period, but the incorporation of this parameter into the data does not significantly increase the variance of the subsequent power distribution.

This lower variability of wave power can be considered a major advantage in terms of providing a more reliable source of energy, but also because it allows the generator used in any wave energy technology to be more precisely rated, resulting in a higher degree of electrical efficiency. This has major implications for the often discussed concept of capacity factor, which is, therefore, implicitly much higher for wave energy.

2. Fitting wind and wave data to a mathematically tractable theoretical distribution

Wind speed and wave height data have traditionally been fitted to Weibull or Rayleigh distributions, the latter being a special case of the former with shape parameter equal to two. In general, a Weibull distribution represents the reality of these distributions quite well, but it is not particularly easy to work with mathematically. Rayleigh distributions, while simpler, often do not do a good job of modeling the real data. In cases of bi-modal distributions, as can occur with wind data at certain locations, neither the Weibull nor Rayleigh distributions are at all accurate.

One of the motivations for this paper is to represent wind and wave data accurately, while also providing a suitable framework for the simple mathematical determination of higher order moments, which is necessary for calculating the variance of the corresponding distributions of power. The distribution that was chosen bears close similarity to both the Weibull distribution and the gamma function. Its extra degree of freedom allows for a much better fit than is often the case with a Rayleigh distribution and, qualitatively, it competes on a par with the Weibull distribution for accuracy. It is also very simple to work with mathematically.

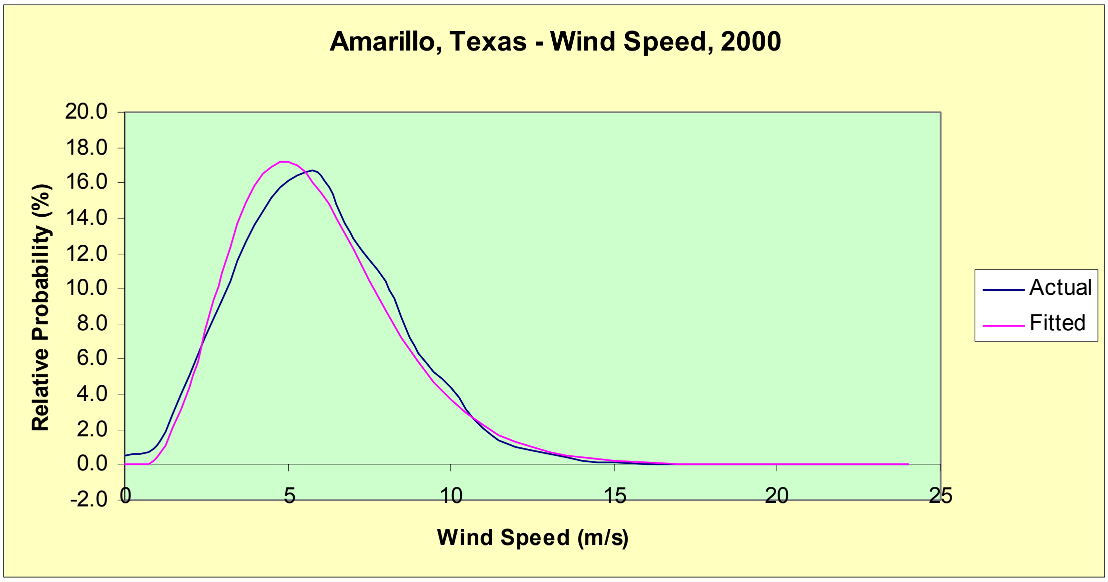
The form of this distribution is

$$f(x) = \alpha \; x^{n-1} \; e^{-\mu x} \qquad (1)$$

The three parameters α, n, and μ are chosen so that the curve optimally fits the real distribution. This is attained via a least squares minimization, constrained by the requirement to retain the same weighted mean and a sum of probabilities equal to unity.

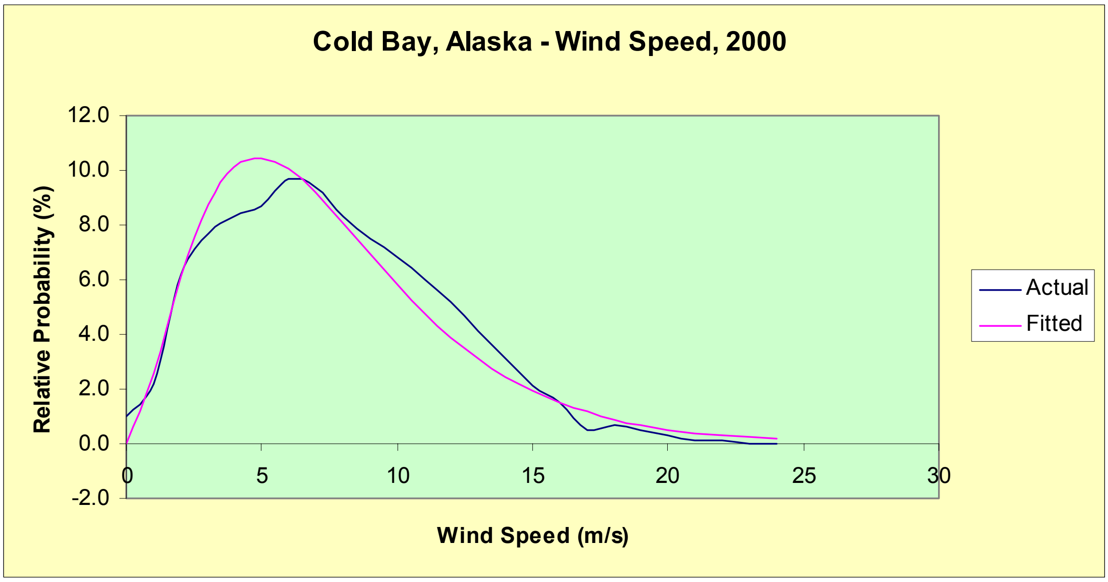
Several wind speed and wave height data sets were analyzed and the curve from equation (1) was optimally fitted to the distributions. These included wind speed data, taken at 10 metres height throughout the year 2000, at Amarillo, Texas, USA; Cold Bay, Alaska; Dodge City, Kansas; and Sandberg, California; and wave height data, recorded approximately one kilometre off-shore in 22 metres of water depth at Port Kembla, Australia from 1987 to 1997; five kilometres off shore in 100m of water at Mokapu Point, Hawaii, during 1999-2000; three kilometres offshore in 50m of water at Bilbao, Spain, from 1985 to 1999; and approximately three kilometres off-shore in 50 metres of water depth off Vancouver Island, Canada during 1993-95. The resulting curves are displayed in Figures 1 – 8, where both the real data curve (actual) and the curve of best fit (fitted) are illustrated in each figure, along with the value of the "best fit" parameters α, n, and μ.

As can be seen, the curve fits for the Amarillo, Dodge City, and Sandberg wind speed data and the Port Kembla and Hawaiian wave height data are exceptionally good. The other three data sets are represented quite well also, even though they were chosen because the real data curves were more unusually shaped, and it was expected it would be difficult to fit theoretical curves to these distributions. In general, the curve fits across the whole data set are very good. The conclusion is that the use of the statistical distribution of equation (1) in the representation of wind speed and wave height distributions is well justified.



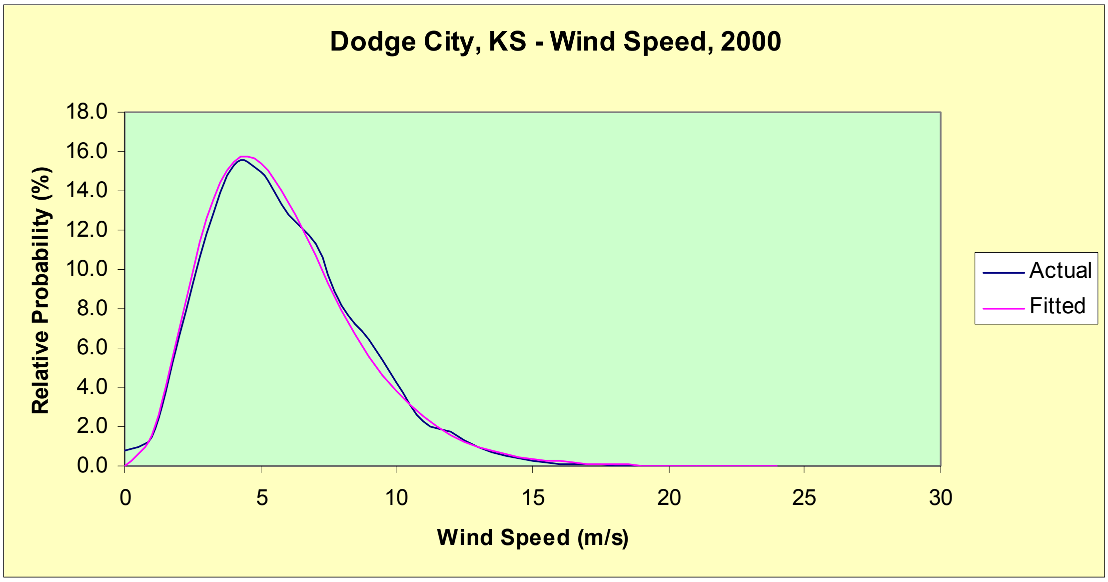
Source: www.stanford.edu

Figure 1 $\alpha = 1.199,\ n = 5.573,\ \mu = 0.939$



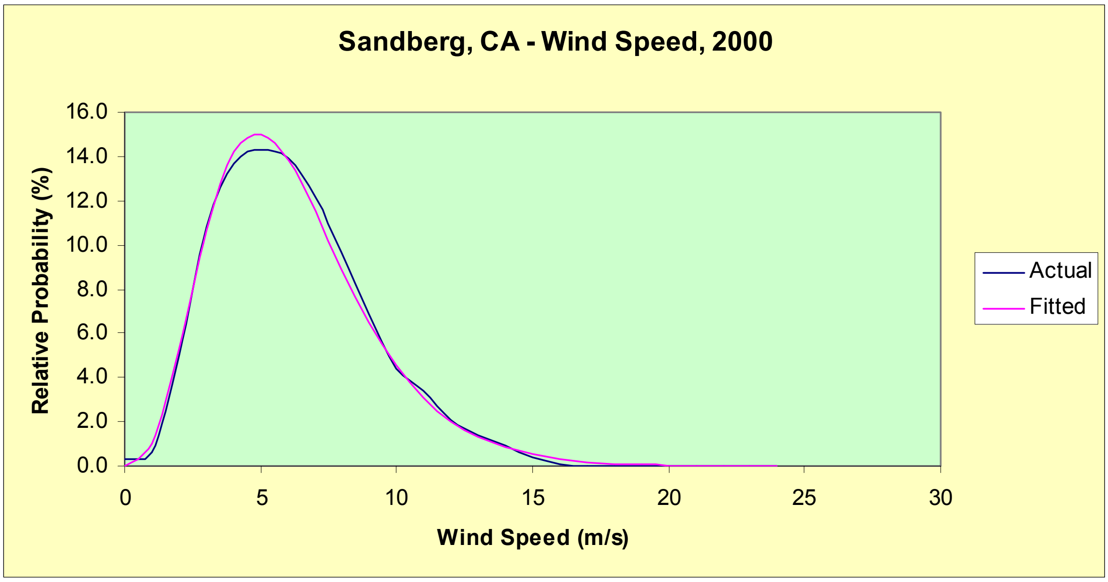
Source: www.stanford.edu

Figure 2 $\alpha = 3.717,\ n = 2.774,\ \mu = 0.364$



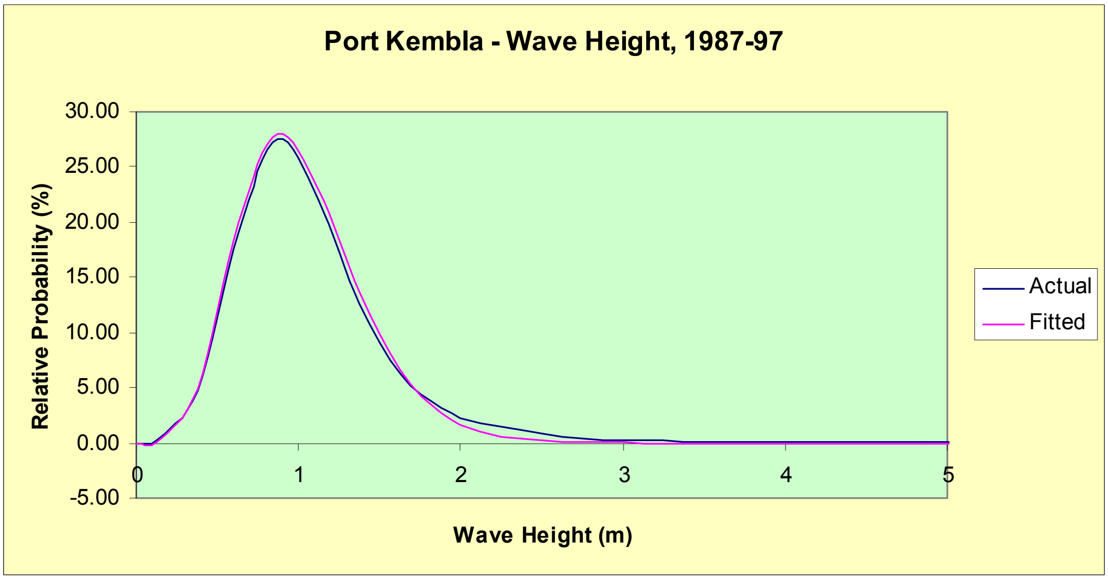
Source: www.stanford.edu

Figure 3 $\alpha = 3.16,\ n = 4.25\ ,\ \mu\ = 0.73$



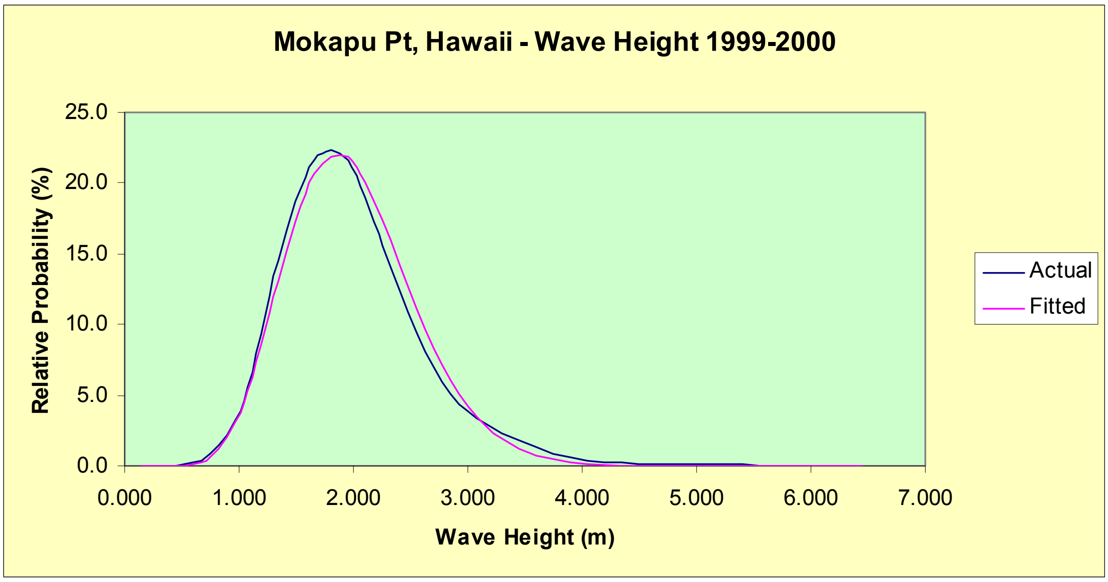
Source: www.stanford.edu

Figure 4 $\alpha = 2.00,\ n = 4.50\ ,\ \mu\ = 0.72$



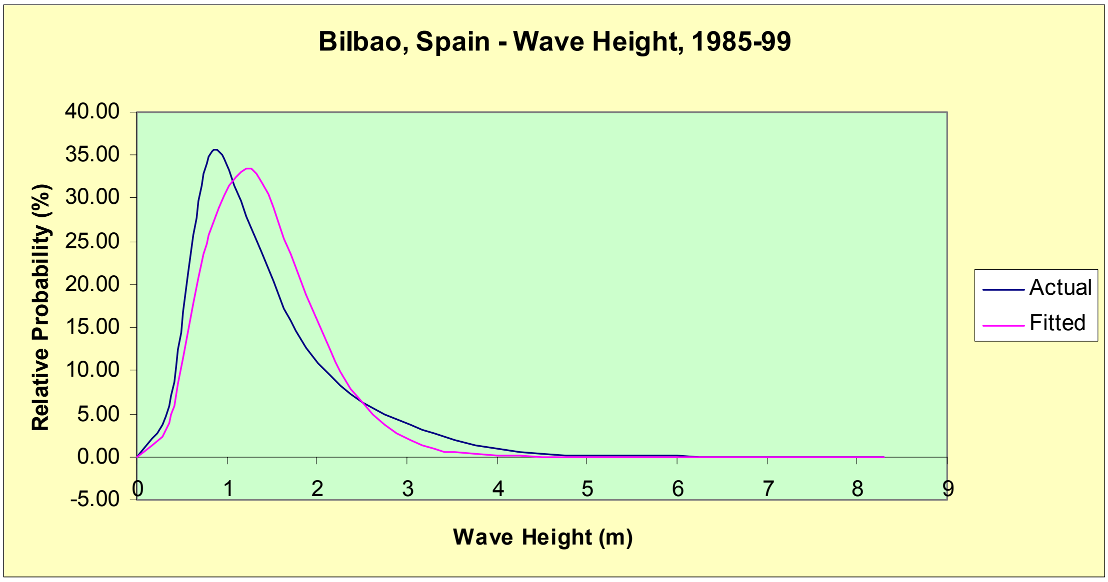
Source: Port Kembla Port Corporation

Figure 5 $\alpha = 31239$, $n = 7.582$, $\mu = 7.066$



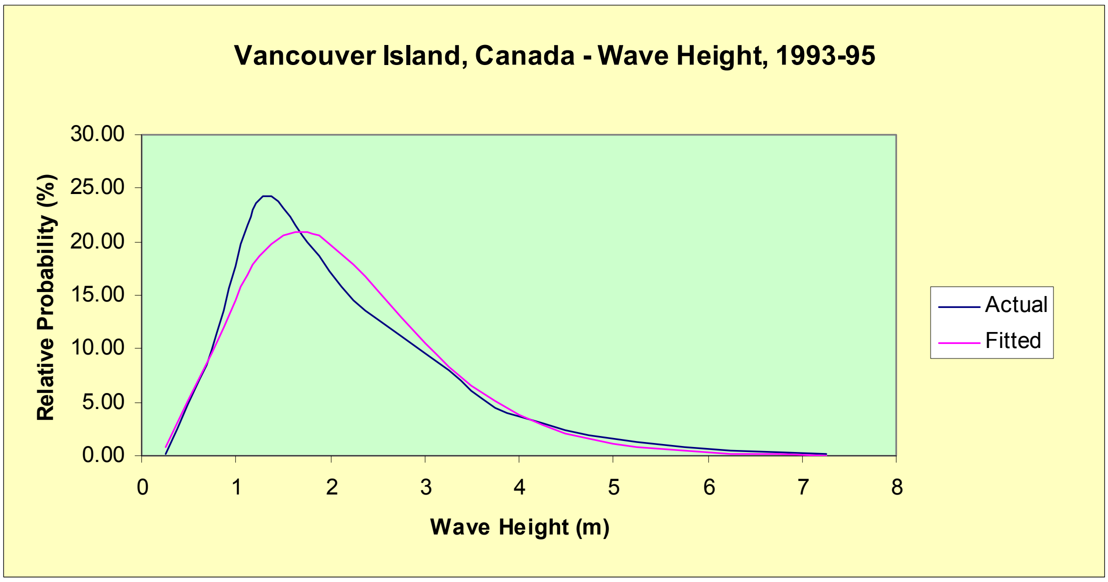
Source: BRG Group

Figure 6 $\alpha = 2310$, $n = 13.03$, $\mu = 6.51$



Source: Bilbao Port Authority

Figure 7 $\alpha = 1437,\ n = 5.40,\ \mu = 3.79$



Source: BC Hydro

Figure 8 $\alpha = 102.63,\ n = 4.201,\ \mu = 1.932$

3. Calculation of moments

The first moment of a distribution is its mean or average. This is then used to calculate the standard deviation of the distribution. However, in order to calculate the standard deviation of the square of a distribution (power from wave height) or cube of a distribution (power from wind speed), it is necessary to calculate higher order moments.

In general, given the form of equation (1), the m[th] moment of the distribution is

$$\overline{x^m} = \frac{\int_0^\infty x^m\, f(x)\, dx}{\int_0^\infty f(x)\, dx} \tag{2}$$

$$\therefore \qquad \overline{x^m} = \frac{\alpha \int_0^\infty x^{m+n-1} e^{-\mu x}\, dx}{\alpha \int_0^\infty x^{n-1} e^{-\mu x}\, dx} \tag{3}$$

Therefore, upon evaluating these integrals, the m[th] moment of such a distribution is

$$\overline{x^m} = \frac{(m+n-1)!}{(n-1)!\,\mu^m} \tag{4}$$

The corresponding variance is

$$\sigma_{x^m}^2 = \frac{\int_0^\infty (x^m - \overline{x^m})^2\, f(x) dx}{\int_0^\infty f(x) dx} \tag{5}$$

$$= \frac{\alpha \int_0^\infty \{x^{2m+n-1} - 2x^{m+n-1}\,\overline{x^m} + x^{n-1}\,\overline{x^m}^2\} e^{-\mu x}\, dx}{\alpha \int_0^\infty x^{n-1}\, e^{-\mu x}\, dx} \tag{6}$$

$$= \frac{(2m+n-1)!(n-1)! - (m+n-1)!^2}{(n-1)!^2\, \mu^{2m}} \tag{7}$$

Therefore, the standard deviation of the distribution (the square root of the variance) is

$$\sigma_{x^m} = \frac{\sqrt{(2m+n-1)!(n-1)! - (m+n-1)!^2}}{(n-1)!\,\mu^m} \tag{8}$$

When comparing the variability of different distributions, the important measure is the relative variance or relative deviation. The relative deviation is defined as the ratio of the standard deviation to the mean of the distribution. Therefore, the relative deviation of the mth moment is

$$\psi_m = \frac{\sigma_{x^m}}{\overline{x^m}} = \sqrt{\frac{(2m+n-1)!(n-1)!}{(m+n-1)!^2} - 1} \tag{9}$$

Interestingly, the relative deviation does not depend at all on the fitted distribution parameters α and μ. Instead, the relative variability of a distribution is entirely dependent on the fitted parameter n and the degree of its moment m.

4. Comparing the variability of wind speeds and wave heights

Equation (9) provides a simple and convenient way to compare the variability of any two distributions fitted to a curve as defined in equation (1). Using equation (9), we see that the means (first moments), standard deviations, and relative deviations of the eight distributions examined above are as follows (note, the wind data are expressed in metres per second, while the wave data are in metres):

Table 1

	Mean	Standard Deviation	Relative Deviation
Amarillo, TX	5.932	2.513	0.424
Cold Bay, AK	7.618	4.574	0.600
Dodge City, KS	5.821	2.822	0.485
Sandberg, CA	6.222	2.933	0.471
Port Kembla, AUS	1.073	0.390	0.363
Mokapu Point, HA	2.002	0.555	0.277
Bilbao, Spain	1.427	0.614	0.430
Vancouver Island, CAN	2.174	1.061	0.488

It does make sense to compare actual means and standard deviations of these distributions, since this entails comparing the units of wind speeds with those of wave heights. However, comparison of the relative deviations is non-dimensional and, therefore, perfectly acceptable. The table above indicates there is no clear-cut statistically significant difference, in this small sample set, between wind speeds and wave heights in terms of their variability. The Hawaiian and Port Kembla wave heights are the least variable of these eight data sets, followed by the Amarillo wind speeds and Bilbao wave heights. The most variable are the Vancouver wave heights and Cold Bay wind speeds. In fact, the latter two were specifically chosen because, in the interests of having a wide range of cases, they were expected to be more variable than the others. This is indeed the case.

In general though, from this data, one cannot say that wind speeds are more or less variable than wave heights. They are of the same order of magnitude. A much more thorough analysis of many data sets of wind speeds and wave heights from around the world would be needed to justify a distinction between the two. Therefore, for the purposes of the conjecture of this paper, let it be assumed there is no intrinsic variability difference between wind speeds and wave heights. This assumption is consistent with the results of the small sample set of Table 1.

5. Comparing the variability of wind power and wave power

While there may not be any innate difference in the relative variability of wind speeds and wave heights, the same cannot be said of the corresponding distribution of power for the two phenomena. Under the assumption that the raw wind speed and wave height distributions have the same degree of variability, there must then be greater variability in the cube of such distributions compared to the square of these distributions.

The power in swept area A for wind with a velocity u and an air density ρ_{air} , is given by

$$P_{wind} = 0.5\,\rho_{air}\,A\,u^3 \tag{10}$$

Correspondingly, the power per length of wave crest L in a water wave of height H, with a group velocity of c_g and a water density ρ_{water} , acting under a gravitational acceleration of g, is

$$\mathrm{P_{wave}} = \frac{1}{8}\rho_{\mathrm{water}}\,\mathrm{g\,H^2\,L\,c_g} \tag{11}$$

In both cases, a suitable choice of area A and length L will lead to the co-efficients of u^3 and H^2 being unity, without altering the relative deviation in any way (since any co-efficient of these terms will apply equivalently to the mean and standard deviation, thus canceling in the calculation of the relative deviation).

Therefore, the relative deviations of the power in the wind and the waves can be evaluated by simply considering the relative deviations of the cube of wind speed and square of wave height (as mentioned previously, wave power depends also on the wave period – however, this parameter has a much lesser effect on the level of power, and is highly unlikely to alter the broad outcome of these results to any great degree - this conjecture will be confirmed later in this paper). The cube of wind speed and square of wave height can easily be evaluated via the higher moments of these variables.

From equation (9), the relative deviation of a distribution of the cube of wind speed (the power in the wind), with distribution parameter *n,* is given by inserting *m* = 3. That is,

$$\psi_3 = \sqrt{\frac{(n+5)!(n-1)!}{(n+2)!^2} - 1} \tag{12}$$

Likewise, the relative deviation of a distribution of the square of wave height (the power in a wave), with a distribution parameter *n*, is given by inserting *m* = 2 into equation (9). That is,

$$\psi_2 = \sqrt{\frac{(n+3)!(n-1)!}{(n+1)!^2} - 1} \tag{13}$$

Using these expressions, and making use of the gamma function to calculate non-integer factorials, the relative deviation of the power in the resource, for each of the distributions illustrated in Table 1, is displayed in Table 2.

Table 2

	Relative Deviation
Amarillo, TX	1.459
Cold Bay, AK	2.255
Dodge City, KS	1.722
Sandberg, CA	1.663
Port Kembla, AUS	0.747
Mokapu Point, HA	0.564
Bilbao, Spain	0.894
Vancouver Island, CAN	1.022

Note, whereas there is no clear distinction between the variability of wind speeds and wave heights, as evidenced in Table 1 (where, in fact, wave heights for Vancouver Island and Bilbao were relatively more variable than wind speeds at Amarillo), there does appear to be a distinction between the variability of the power contained in these respective phenomena. As evidenced in Table 2, the power in the waves at Vancouver Island and Bilbao is noticeably less variable than the power in the wind at Amarillo. As expected, the variability of the power in the waves in Hawaii and at Port Kembla is very low compared to the others, especially compared to the highly variable power in the wind at Cold Bay.

These facts are illustrated in Figure 9, where both the relative deviations of the wind speed and wave height data are presented graphically, along with the corresponding relative deviations of their associated power levels. The first four sets of bars represent the wind sites, and the second four the wave sites. Once again, whereas it is difficult to differentiate between the variability of wind speeds and wave heights at these sites, there is a clear distinction between the variability of the corresponding power levels at the sites.

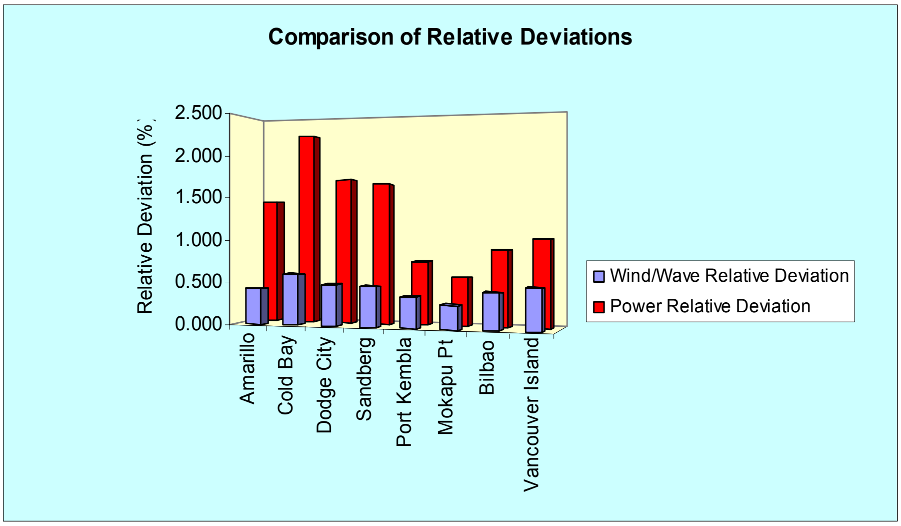


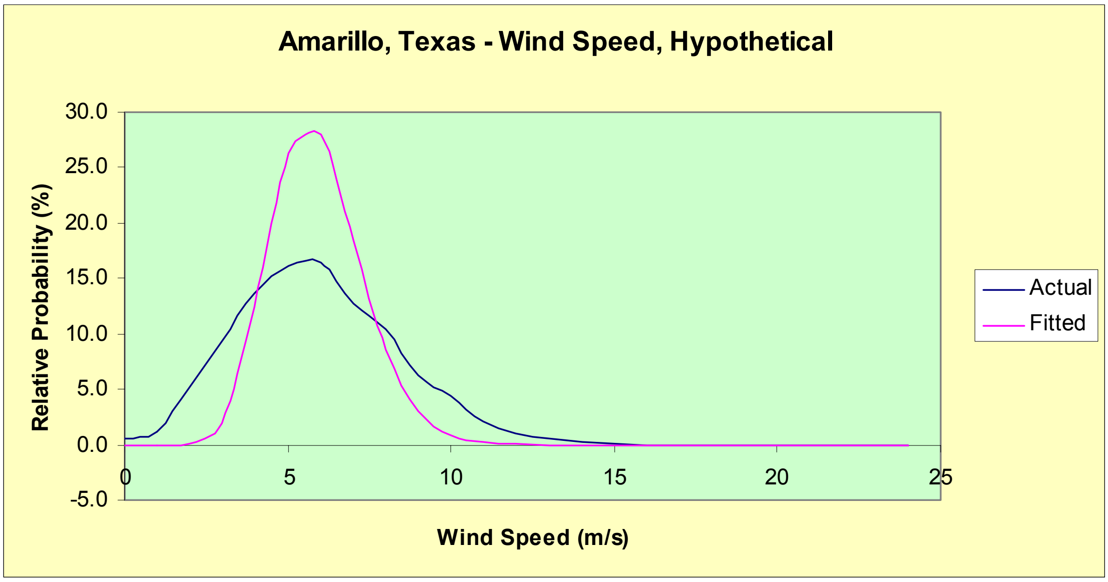
Figure 9

This analysis does not categorically show that the power in the wind at any site around the world will be more variable than the power in the waves at any site. It is entirely possible that there are sites with a very low degree of variability in the wind. However, assuming that the variability of wind speeds and wave heights is generally on a par (although a case can easily be made that there is some added degree of variability in wind speeds compared to wave heights – this hypothesis is not argued in this paper), it is therefore true, as shown above, that there is a higher probability the power in the wind at any particular site will be more variable than the power in the waves at any particular site.

6. A hypothetical wind distribution with an equivalent variability as a wave distribution

For the purpose of quantifying the level of consistency required in a distribution of wind speeds, so that the variability of its corresponding power is on a par with the variability of the power in a typical wave height distribution, the following analysis was performed. The parameters for the fitted distribution for Amarillo, TX, were adjusted so that the relative deviation of the power of this artificial distribution was equal to the actual relative deviation of the power in the real waves at Port Kembla. The resulting wind speed distribution for this new “less variable” artificial curve is displayed in Figure 10, along with the real distribution for Amarillo. The means of the two distributions are equal, but the variability of the wind speeds, and particularly the variability of the power in the wind, is quite different.

Another way to view this result is that any wind speed distribution with the same mean wind speed as Amarillo, would need to have wind speeds banded almost entirely within a range of 3 m/s to 10 m/s in order for the variability in the power of the wind at that site to be equal to the actual variability of the power in the waves at Port Kembla. Such a narrow band of wind speeds is extremely uncommon.



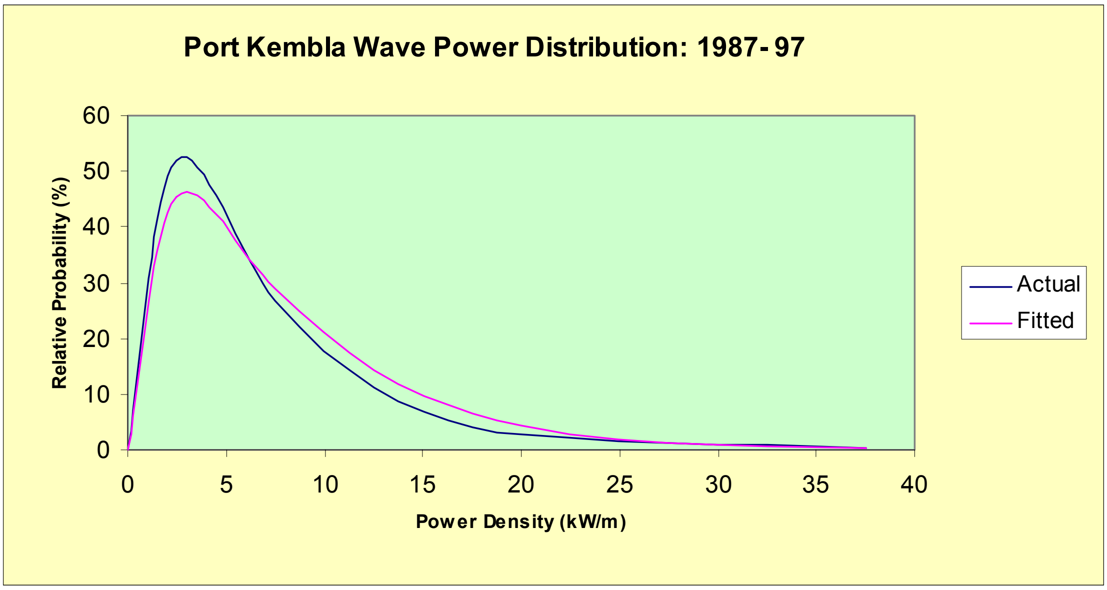
Source: www.stanford.edu

Figure 10 $\alpha = 0.00015,\ n = 17.89,\ \mu = 3.017$

7. Incorporating wave period and depth

The analysis of wave power variability so far has been restricted to wave height as the sole independent determinant. As stated earlier in this paper, wave power levels also depend, to a lesser degree, on wave period and depth. As a means of quantifying what these parameters mean to the variability of wave power, the actual power distribution for waves at Port Kembla was constructed, using the full dispersion relation for finite depth surface gravity waves, to dictate the true wavelength and, hence, the group velocity, for all combinations of wave heights and periods at the site in the 22 metre water depth.

This resulting distribution of wave power is illustrated in Figure 11, for the real case and that of a curve fitted to the data via the distribution defined in equation (1). Both the actual and fitted distributions have a mean of 7.73 kW per metre of wave crest and a relative deviation of 0.859. This relative deviation incorporates variability due to all the wave power input variables. The value of 0.859 varies from the relative deviation for the Port Kembla wave power distribution of 0.747 (calculated using the wave height distribution only) by only 15%. This additional 15% of variability can be attributed to the effect of the variability of the wave period and how it interacts with water depth to affect wave length. As can be seen, the effect is not pronounced, and the overall relative deviation of 0.859 is still very low compared to that of the power distributions calculated from the wind speed data. This result is consistent with the claim that wave power, in general, is less variable than wind power, regardless of the subtleties and approximations in how it is calculated.



Source: Port Kembla Port Corporation

Figure 11 $\alpha = 102269$, $n = 1.429$, $\mu = 0.185$

8. Capacity Factors

The capacity factor of a renewable power plant is defined as the ratio of the average annual power output to the peak power output. Power plants which depend on fuel (coal, gas, biomass etc.) have a theoretical capacity factor of nearly 100%, as they can, in theory, be run at the peak rated capacity all the time (apart from periods of maintenance).

However, a renewable power plant, relying on a variable input source, must innately have a rated peak capacity above the average annual capacity. This ratio will be higher the less variable the power source. Therefore, if wave energy is less variable than wind energy, it would be expected that wave power plants have higher capacity factors. Capacity factors for wind power have been established in the range of 25% to 35%. No wave power plant has operated long enough to establish a corresponding bench mark capacity factor for wave power.

The capacity factor for a particular site, derived from the wind speed or wave height data, can be readily calculated using the relative deviation of that distribution. A commonly acceptable practice is to rate a renewable power generator such that, on 10% of occasions, it will be running above this rated capacity. This allows for a balance between maximizing output and minimizing the electrical losses incurred from running the generator at below its rated capacity for much of the time.

The 10% cut-off point was calculated for each of the eight distributions in Table 1, and expressed in terms of how many standard deviations it is above its mean. For a normal or Gaussian distribution, this point γ_r, at which 90% of the distribution lies below and 10% above, is about 1.3, meaning it lies 1.3 standard deviations above the mean. However, wind and wave distributions are not Gaussian. The sites examined in this paper display values for γ_r ranging between 1.25 and 1.48. These are illustrated in Table 3.

The rated power of a generator P_r, defined according to the "10% rule", with mean power $\overline{P}$ and operating at a site with a relative deviation of its power source of ψ_m and rated cut-off point of γ_r, will therefore be defined as

$$P_r = \overline{P}(1+\gamma_r \psi_m) \tag{14}$$

Therefore, the capacity factor C_f, which is the ratio of $\overline{P}$ to P_r, is

$$C_f = \frac{1}{1+\gamma_r \psi_m} \tag{15}$$

The capacity factor for the eight data sets is also illustrated in Table 3 and shown graphically in Figure 12. Capacity factors for wind are generally in the range of 25-35%. This is consistent with the results of this analysis, with Amarillo exhibiting a capacity factor of 34% and Cold Bay 25%, with Dodge City and Sandberg both around 30%. Clearly, the capacity factor for a wind site varies markedly, depending on the particular distribution of wind speeds.

This does not appear to be so much the case with wave energy. Port Kembla, Bilbao, and Vancouver Island all have capacity factors of 44%. The one anomaly is Mokapu Point in Hawaii, with an implied capacity factor of 54%. This is almost certainly due to the regularity of wave swell generated by the trade winds of the North Pacific, and is consistent with the low variability of the wave height data at this location. Such a site appears ideal for wave energy projects.

These results are quite consistent with research conducted previously by the author. This previous research, although quite different in its approach, found that four different wave energy sites – Port Kembla, San Francisco, Hawaii, and Bilbao (Spain) - all exhibited capacity factors in the vicinity of 45%, when calculated according to the "10% rule". In that study the actual Hawaiian result did display the highest capacity factor of the four sites, at over 46%, though still noticeably lower than the 54% expressed in the current analysis for the same site.

Table 3

	γ_r	C_f (%)
Amarillo, TX	1.339	33.9
Cold Bay, AK	1.308	25.3
Dodge City, KS	1.339	30.3
Sandberg, CA	1.288	31.8
Port Kembla, AUS	1.481	44.0
Mokapu Point, HA	1.349	54.4
Bilbao, Spain	1.308	43.7
Vancouver Island, CAN	1.250	43.9

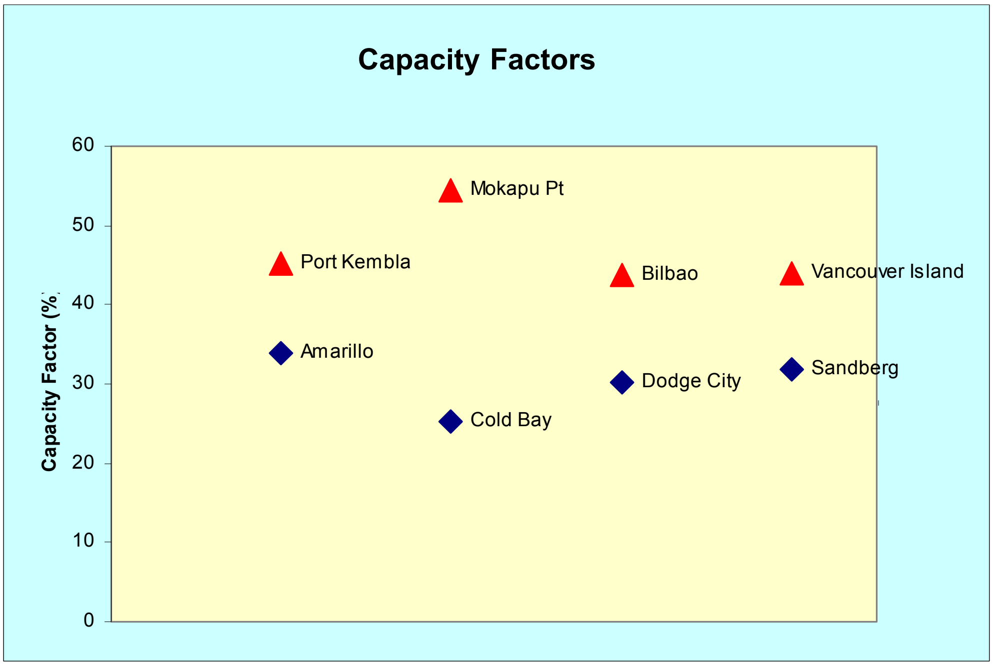


Figure 12 Wind sites – blue diamonds; Wave sites – red triangles

On a more theoretical level, assuming that a typical wind site and a typical wave site have the same natural variability, it is possible to evaluate their respective "typical" capacity factors. For convenience, assume such a typical wave site exhibits a relative deviation of its power of 1 (this is much closer to the relative power deviation of Vancouver Island than Hawaii and Port Kembla, and so, is probably on the high side as indicative of a typical site). Using equation (13) and equating ψ_2 to the value of 1, allows the calculation of the corresponding value of n. In this case, we arrive at a quadratic equation

$$n^2 - 3n - 6 = 0 \tag{16}$$

This equation has a positive solution of $n \approx$ 4.3723. Substituting this value into equation (12) gives $\psi_3 = 1.693$. Taking the average value of γ_r from Table 3 (i.e. 1.3445), and substituting this and the values of ψ_2 and ψ_3 into equation (15), capacity factors of 30.7% for a typical wind site and 42.9 % for a typical wave site are found. Given that capacity factors for wave energy appear to be higher than this value, suggests either that the typical relative deviation of wave energy is actually less than unity (very likely), or that the typical relative deviation of wind speeds is actually greater than that of wave heights (quite likely), or a combination of both.

9. Conclusion

The theoretical distribution, as defined in equation (1), appears to be both a very adequate fit for real wind speed and wave height data, and a simple means of calculating higher order moments and thus power distributions for these phenomena. It also allows for an intuitive understanding of why wave power is less variable than wind power, despite the underlying phenomena being equally variable – that is, due to the reliance on the square of the wave height, compared to the cube of wind speed. Relative deviations for a range of wind and wave sites (covering a wide range of variability within each of these categories) clearly show that wind energy is typically more variable than wave energy. Conversely, a wind site would need to exhibit a very narrow band of wind speeds in order to have an equivalent relative deviation of power as a typical wave site.

Using these theoretical distributions, it is possible to derive a capacity factor, based upon the criterion that a generator is rated such that it will run above its rating on 10% of occasions. These resulting theoretical capacity factors for wind energy are consistent with the 25-35% that is observed with real wind turbines, and the corresponding capacity factors for wave energy are consistent with the 45% derived by the author in a previous study of four sites, based purely on data analysis. The analysis of additional wind and wave data sets would be desirable to further clarify these claims.

ETSU, a unit supporting the UK DTI (Department of Trade and Industry), refers to a term called the declared net capacity (DNC) factor. This incorporates the losses of running the plant, but more importantly, compares the renewable plant's capacity factor to the capacity factor of a conventional power plant, which is claimed to be 70%. The following is an extract from an ETSU publication: "*DNC* is, very broadly, the equivalent capacity of a base-load plant that would produce the same average annual energy output as the renewable energy plant. For wind farms the DNC is calculated by subtracting the on-site electrical power consumption from the installed capacity and multiplying the remainder by 0.43 (i.e. 43%). The reason 0.43 is used, and not the commonly assumed capacity factor of 0.3, is because conventional generating plants also operate at less than their maximum output for much of the time. When DNC factors were originally established, it was assumed that wind farms would operate with a capacity factor of 0.3 and a conventional plant with a capacity factor of 0.7. The DNC was therefore defined as (0.3÷0.7) or 0.43". Using this definition, a wave power plant would have a DNC of around 64% (minus small running losses). In other words, <u>any wave energy plant should average at least 64% of the output of an equivalently rated conventional fossil fuel power plant.</u>

The capacity factors discussed here assume that a renewable technology is able to convert all the power from the source, or at least in some constant proportion across the distribution. In reality, depending on the conversion device, it may be possible to increase the capacity factor by converting more at the low end of the distribution and less at the high end, thus moving the converted average closer to the 10% cut-off of the distribution. This is indeed what is apparent from Energetech's research into the use of retractable parabolic walls in conjunction with an oscillating water column. Capacity factors approaching 60% (and declared net capacity factors approaching 85%) appear possible under this scenario.