Home Assignment 1: Random Number Generation, Power production of a wind turbine and Combined power production of two wind turbines

FMSN50 - Monte Carlo and Empirical Methods for Stochastic Inference

Victoria Lagerstedt & Kajsa Hansson Willis

March 10, 2025

Contents

1 Introduction

Renewable energy sources are playing an increasingly important role in worldwide energy creation. Among them, wind power is a key contributor due to its sustainability and technological advancements. However, evaluating the feasibility and efficiency of a wind turbine at a given location requires a detailed statistical analysis of wind speed distributions and power output.

This report uses Monte Carlo methods and empirical statistical techniques to analyze the expected power production of a Vestas V164 9.5 MW wind turbine based on wind conditions from a site in northern Europe. Wind speeds are modeled using a Weibull distribution and are leveraged in an empirically determined power curve. Several variance reduction techniques—including importance sampling, control variates, and antithetic sampling—are implemented to improve the efficiency of Monte Carlo estimations. Further, bivariate distributions are considered.

2 Part 1: Random number generation

2.1 a) Conditional distribution function and density given an interval

In the first part of the assignment, the conditional distribution function

$$F_{X|X\in I}(x) = P(X \leq x|X \in I)$$

and the density function were determined. X is here a real random variable, and I=(a,b) is an interval.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{1}$$

$$P(A \cap B) = \int_a^x f_X(z)\,dz = F_X(x) - F_X(a) \tag{2}$$

$$P(B) = \int_a^b f_X(z)\,dz = F_X(b) - F_X(a) \tag{3}$$

To calculate the conditional distribution function Equation 1, Equation 2 and Equation 3 were used. The conditional distribution density is therefore described by

$$\frac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)}$$

inside of the interval (a,b), zero if x is smaller than a and one if x is larger than b. This is of course because if x is smaller than a none of stochastically determined values can be smaller than that range, and similarly, if x is larger than b then none of the stochastically determined values can be smaller than x, since they are all within the interval.

To calculate the density function, the derivative of the conditional probability function was calculated with equation 4 for a<x<b and zero otherwise.

$$f_{X|X\in I}(x) = \frac{d}{dx}F_{X|X\in I}(x) = \frac{f_X(x)}{F_X(b) - F_X(a)} \tag{4}$$

2.2 b) Inverse function

The inverse conditional distribution function is calculated from the conditional distribution function with Equation 5, where I = (a,b) is an interval.

$$F_{X|X\in I}(x)^{-1} = F_X^{-1}(F_{X|X\in I}(x) * (F_{X|X\in I}(b) - F_{X|X\in I}(a)) + F_{X|X\in I}(a)) \tag{5}$$

This inverse conditional distribution can be used to simulate X conditionally on X in I. To simulate this, one can use inverse transform sampling. The first step is to draw random samples from a uniform distribution U in [0,1] and then find the function T that solves T(U)=X. Since uniform samples this can be done by following equation 6.

$$F_X(x) = P(X \leq x) = P(T(U) \leq x) = (U \leq T^{-1}(x)) = T^{-1}(x) \tag{6}$$

Due to these equalities, T(x) is the same as the inverse conditional distribution of x. To simulate X, uniform random variables are used as input into the inverse conditional distribution.

3 Part 2: Power production of a wind turbine

Here the potential of a wind turbine at the specific location is evaluated. In each of the following parts, an answer was created for each month by running a loop. V is the wind speed and P(V) is the power generated by the power turbine when the wind speed is V.

3.1 a) Confidence interval drawing from the Weibull distribution

In this part of the assignment, the purpose was to create a 95% confidence interval of the expected amount of power generated by the wind turbine using draws from a Weibull distribution. Further, another confidence interval was supposed to be created based on the truncated version described in part 2.

The confidence interval can be determined according to Equation 7, where τ is the expected value σ the standard deviation, N is the sample size, and p is the corresponding p-value from the normal distribution for the degree of confidence interval (1.96 for a two-sided 95% confidence interval).

$$\text{Confidence interval} = [\tau - \frac{\sigma * p}{\sqrt{N}}, \tau + \frac{\sigma * p}{\sqrt{N}}] \tag{7}$$

In the first step of the Matlab code, vectors were created to store the λ and k values for each month, so that vectors could store each month's expected value and confidence interval. For each month, a random Weibull distribution was generated. These values were then inserted into the power curve and the mean was saved as the expected value. Additionally, the standard deviation was calculated for the power curve of the distribution. These values of τ and σ were then inserted into Equation 7 where N was chosen to be 10 000 samples. The results from the standard Monte Carlo can be viewed in Table 1, where the average of each column has been added as a yearly estimate.

Month:	Mean value (W):	Lower bound (W):	Upper bound (W):	Width:
January	4612986	4565966	4710358	144392
February	4162590	4050305	4188583	138278
March	3862859	3711460	3846808	135348
April	2998917	2916247	3040740	124493
May	2896244	2763523	2885496	121973
June	3083294	3064023	3191759	127736
July	2872771	2826996	2949610	122614
August	3060296	3005860	3132857	126997
September	3765101	3748644	3884213	135569
October	4215729	4174607	4316852	142245
November	4611597	4579472	4723601	144129
December	4638101	4571896	4716416	144520
Yearly	3731700	3664916	3798941	134024

Table 1: Estimation of power generation with the expected value, lower and upper bounds of the 95% confidence interval including width for each month and the average as a yearly estimate using standard Monte Carlo

After this, the truncated version was calculated. For each month N random uniform samples were drawn and used in the inverse conditional probability function to create random Weibull samples. This instead of using Matlab's Weibull random generator. The conditional probability function of a Weibull distribution can be seen in equation 8. The inverse probability function of a Weibull distribution was calculated from equations in part 2. The result is seen in Equation 9, where k(i) means the k-value for the month with the index i.

Month:	Mean(W):	Lower bound (W):	Upper bound (W):	Width of confidence interval:
January	4682473	4621226	4743721	122495
February	4166130	4106434	4225826	119392
March	3857431	3800665	3914197	113532
April	3000061	2949595	3050526	100931
May	2873465	2824376	2922554	98178
June	3071130	3020197	3122062	101865
July	2919405	2870259	2968551	98292
August	3111477	3060315	3162639	102324
September	3767875	3711423	3824327	112904
October	4194426	4135135	4253717	118582
November	4691586	4630728	4752444	121716
December	4708731	4647830	4769633	121803
Yearly	**3753682**	**3698181**	**3809183**	**122495**

Table 2: Estimation of power generation with the expected value, lower and upper bounds of the 95% confidence interval including width for each month and the average as a yearly estimate using truncated Monte Carlo

$$F_x = 1 - \exp\left(-\left(\frac{x}{\lambda(i)}\right)^{k(i)}\right) \tag{8}$$

$$F_{X|X\in I}(x)^{-1} = \lambda(i) * (-ln(e^{-(3.5/\lambda(i))^{k(i)}} - (e^{-(3.5/\lambda(i))^{k(i)}} - e^{-(25/\lambda(i))^{k(i)}}) * F_{X|X\in I}(x)))^{1/k(i)} \tag{9}$$

Since the interval, I, represented in the truncated version was [3.5,25] m/s, these were the values used in Equation 9. The obtained values were inserted into the power curve to estimate the generated power from the distribution. This value was then multiplied by the likelihood of the value being in the interval, i.e. $F(b) - F(a)$, where b=25 and a=3.5, to account for the truncation in the estimate of the power generation for all intervals. Extracting the mean generated the expected value. Further, the standard deviation was extracted and the values were, as previously, inserted into Equation 7 to create the confidence interval. The results using the truncated Monte Carlo can be observed in Table 2.

The expected amount of generated power from the truncated version is slightly larger than from the standard Monte Carlo, but the difference is less than 0.6%. Further, the confidence interval is, on average, noticeably narrower. It is almost 10% larger for the standard Monte Carlo, which is a significant difference. This indicates that the truncation contributes considerable variance reduction. It is also noteworthy that the confidence interval is smaller for every single month.

3.2 b) Using wind as a control variate

One method that decreases the variance of a Monte Carlo simulation is to use control variates. In this case, wind was used as the control variate. To do this equation 10 was used, where ϕ(x) is the power of the simulated wind, Y is simulated wind and m is the mean of Y. It is very important that m is calculated analytically and not estimated for the integrity of the control variate estimation. Beta is the covariance between ϕ(x) and Y divided by the variance of Y, times -1. The 95% confidence interval was calculated using equation 7.

$$Z = \phi(X) + \beta(Y - m) \tag{10}$$

To calculate the value of m, Equation 18 was used where $m = E(V)$, so the m in the equation is equal to one.

The results of the estimation with wind as a control variate can be observed in Table 3. The width of the confidence intervals has decreased significantly compared to the standard Monte Carlo and the truncated version, which shows that using the wind as a control variate is variance reducing. The reason for this is that information about errors in the wind distribution is used to decrease the errors in the power estimation.

Month:	Mean(W):	Lower bound (W):	Upper bound (W):	Width of confidence interval:
January	4683508	4651364	4715653	64289
February	4123453	4094182	4152725	58543
March	3852529	3824811	3880246	55435
April	3006024	2976443	3035606	59163
May	2896244	2866585	2925903	59318
June	3084745	3055548	3113942	58394
July	2865064	2834953	2895175	60222
August	3103811	3075355	3132267	56912
September	3760736	3733355	3788116	54761
October	4236382	4207296	4265467	58171
November	4622718	4588891	4656545	67654
December	4640866	4607938	4673794	65856
Yearly	**3739673**	**3709726**	**3769619**	**64289**

Table 3: Estimation of power generation with the expected value, lower and upper bounds of the 95% confidence interval including width for each month and the average as a yearly estimate with wind as a control variate

3.3 c) Using importance sampling

In this part, the purpose was to use importance sampling to construct a 95% confidence interval of the power of the wind turbine. To do so, an importance sampling function, g, must first be determined from an instrumental distribution so that Equation 11 can be applied.

$$\mu = \int \phi(x)f(x)dx = \int \phi(x)\frac{f(x)}{g(x)}g(x)dx \tag{11}$$

To do so, the function $\phi(x)f(x)$, where $\phi(x) = P(v)$, was plotted and then compared to other distributions that could be appropriate instrumental distributions, as can be seen in Figure 1. The function was also evaluated analytically.

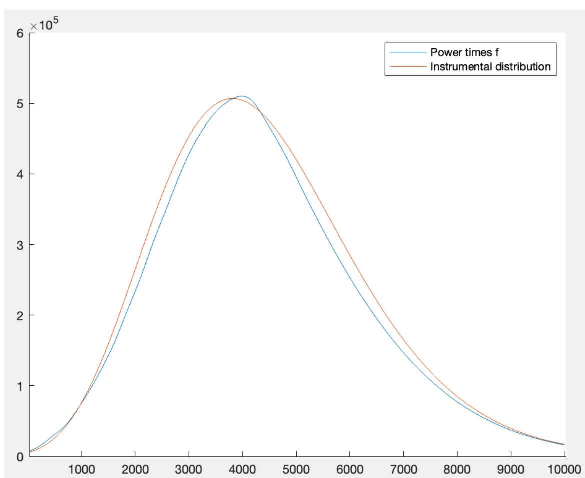

Figure 1: The instrumental distribution and $f(x)\phi(x)$ plotted against each other for January

Further, the ratio between $\phi(x)f(x)$ and g(x), Equation 12, was plotted in Figure 2. If the instrumental distribution is proper, this ratio should be as constant as possible in order to reduce variance as much as possible and shrink the confidence interval.

$$r = \frac{\phi(x)f(x)}{g(x)} \tag{12}$$

The importance sampling that is plotted and was deemed satisfactory is g = 5000000*gampdf(x,10,1.3), which is a Matlab function of the gamma distribution where x is the winds used as input that follow

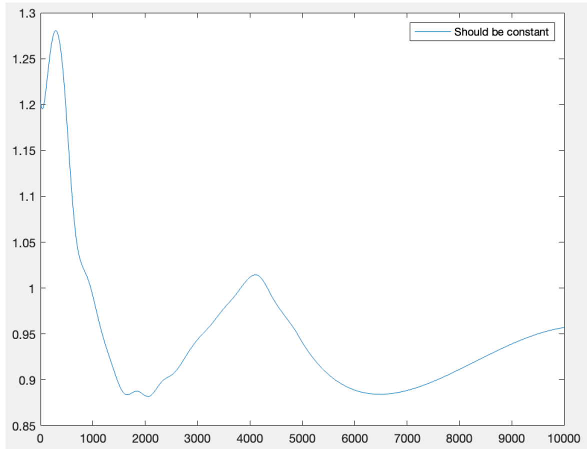

Figure 2: Plot of the ratio described in Equation 12 for January

the Weibull distribution and 10 and 1.3 are the shape and scale parameters. The scalar was chosen to enable visualization of the plots, but it was not necessary for further calculations as it would cancel itself out.

The gamma distribution was deemed as an option because of its similar structure to the power curve, and with proper scaling and choice of parameters this g function was chosen as it appeared very similar to the power curve and the ratio appeared to be quite constant, and it satisfied the requirement such that $g(x) = 0 => \phi(x)f(x) = 0$. The parameters for the gamma distribution that were the best fit were 10 as the shape parameter and 1.3 as the scale parameter.

When the importance sampling had been determined, a loop was created so that the following process was done for each month separately with the appropriate lambda and k values in the Weibull process. These values were gathered into a vector and displayed in Table 4.

First, Weibull probability density functions were constructed for the month and samples of the gamma distribution were extracted. The target function was then constructed with the determined values and by taking the mean, the expected value was extracted. Then the standard deviation was calculated and the confidence interval was determined according to Equation 7.

Month	Mean (W)	Lower bound (W)	Upper bound (W)	Width of confidence interval
January	4658503	4649132	4667873	18 741
February	4137374	4120199	4154549	34 350
March	3836589	3814116	3859063	44 947
April	3025950	2996370	3055530	59 160
May	2854193	2823257	2885130	61 873
June	3072233	3043609	3100856	57 247
July	2853416	2822449	2884382	50 933
August	3090459	3062005	3118913	56 908
September	3745987	3722510	3769464	46 954
October	4215118	4203637	4226599	22 962
November	4655079	4645239	4664919	19 680
December	4662216	4652935	4671497	18 562
Yearly	3733926	3712954	3754897	41 943

Table 4: Estimation of the power with the expected value, lower and upper bounds of the 95% confidence interval including width for each month and the average as a yearly estimate using importance sampling

When comparing the width of the confidence intervals with the previous estimation methods, it is apparent that the average width is significantly lower than with any other method. With the importance sampling the average width is only about 42 000 W, compared to values between 62 000 and 134 000 for the other methods. This is a very large variance reduction, so it is clearly a better variance reduction technique than truncation or using control variates in this situation. It is, however, interesting to note that the width of the confidence intervals for the importance sampling method varies more between the months than the other methods do. The width of the confidence interval for the method using wind as a control variate is also significantly lower than for the standard Monte Carlo or the truncated version, which shows that that method also can accomplish significant variance reduction. It is important to note that the largest reduction in variance is recorded for the months with the highest wind speed. This indicates that the importance distribution is only optimal for the windy months and the variance for the not so windy months is around the same as when using control variates.

It is also interesting to observe the expected values between the methods, which, excluding the truncated version that is expected to have a higher value, are very similar to each other, especially the standard Monte Carlo and the control variate method. This gives credibility to the fact that importance sampling is a very good variance reduction method that has very good mean precision and a compact confidence interval.

3.4 d) Using antithetic sampling

Another variance reducing method is antithetic sampling. Antithetic sampling involves generating two sets of samples with negative correlation. The power curve is divided into two parts, with one being monotonously increasing and the other being constant. Due to this, it was concluded to do two different antithetic samplings. For the monotonously increasing part over the interval (3.5, 14), the wind was sampled from a Weibull distribution. First, half of the samples were sampled from a Weibull distribution, using the truncated version with the inverse conditional probability function and inserting uniform samples shown in Part 3.1, with the scale and shape corresponding to the month. The other half of the samples for the first part were generated by inserting 1 minus the uniform samples used in the inverse conditional probability distribution. Both sampling methods can be seen in Equation 13. The correlation between the two halves of the samples for the first part was negative, more exact -0.9971. For the constant part over the interval (14, 25), no samples were drawn. Since it was constant it was enough to calculate the probability for the wind speed to be between 14 and 25 m/s times the maximum effect, since the wind turbine generated maximum effect in this interval. The probability for the wind to be between 14 and 25 m/s was determined through the calculation: $F(25) - F(14)$, where F(x) is the conditional probability distribution for a Weibull distribution.

$$\begin{aligned} V &= F_V^{-1}(U), \\ \tilde{V} &= F_V^{-1}(1-U) \end{aligned} \tag{13}$$

The mean of the increasing part was calculated by taking the mean power of the first half of the samples plus the mean power of the other half of the samples divided by two. For the constant part, the mean was equal to the maximum effect times the probability that the wind speed would be in the 14-25 m/s interval. Then the mean of the two part's mean was calculated by taking the mean of the increasing part times the probability of the wind speed to be between 3.5 and 14 m/s, plus the mean of the constant part. If the wind speed is outside these intervals it generates no power and, thus, will not affect the mean. The variance of the mean was calculated using Equation 14 and a confidence interval of 95% was created using Equation 7. The variance for the constant part was set to zero, as no sampling was performed for that part.

$$\text{var} = \frac{1}{4} * (\text{var}[P(X_1)] + \text{var}[P(X_2)]) + \frac{1}{2} * \text{cov}(P(X_1), P(X_2)) \tag{14}$$

The expected values as well as the confidence interval for the expected power generated by the wind turbine using antithetic sampling can be observed in Table 5. The mean values are around the same as for the other methods for each month. What differs more is the variance, which is expected since antithetic sampling is used to decrease the variance. Since the correlation between the variables in the increasing part was almost -1, this means a large reduction in variance which can also be seen in the results.

Months:	Expected value (W):	Lower bound (W):	Upper bound (W):	Width:
January	4654383	4653955	4654811	856
February	4143094	4142672	4143516	844
March	3845069	3844645	3845492	847
April	3008532	3008134	3008930	796
May	2844602	2844218	2844987	769
June	3089052	3088652	3089453	801
July	2864063	2863672	2864453	781
August	3080135	3079736	3080535	799
September	3746895	3746482	3747308	826
October	4217009	4216587	4217431	844
November	4667183	4666748	4667617	869
December	4656298	4655868	4656728	860
Yearly	**3734692**	**3734280**	**3735105**	**824**

Table 5: Estimation of the power with the expected value, lower and upper bounds of the 95% confidence interval including width for each month and the average as a yearly estimate using antithetic sampling

3.5 e) The probability of the turbine delivering power

In this part of the assignment, the purpose was to estimate the probability that the turbine delivers power $P(P(V) > 0$. As it had been noted that the wind turbine only provides power when the wind speeds are between 3.5 and 25 m/s, the relationship that can be seen in Equation 15 was leveraged where F(x) is described by Equation 8.

$$P(P(V) > 0) = P(3.5 \leq V \leq 25) = F(25) - F(3.5) \tag{15}$$

This relationship was applied to each month and the probabilities of the turbine generating power for each month can be observed in Table 6. It can be observed that the likelihood is larger for the winter months.

Month	Probability of generating power (%)
January	89.3
February	87.7
March	86.5
April	81.2
May	80.4
June	81.6
July	80.4
August	81.6
September	86.2
October	86.8
November	89.3
December	89.3

Table 6: Estimation of the probability that the wind turbine generates power for each month

3.6 f) The average ratio of the actual wind turbine output and the total wind power

This part of the assignment had the objective to create a 95% confidence interval for the average ratio of actual wind turbine output to total wind power, which is calculated with Equation 16. $EP_{\text{tot}}(V)$ is calculated from Equation 17.

$$\frac{EP(V)}{EP_{\text{tot}}(V)} \tag{16}$$

$$P_{\text{tot}}(v) = \frac{1}{2}\rho\pi\frac{d^2}{4}v^3 \tag{17}$$

In Matlab, the $EP_{\text{tot}}(V)$ was first calculated using Equation 18 to determine the expected wind speeds for each month and inserting this into Equation 17.

$$E[V^m] = \Gamma\left(1 + \frac{m}{k}\right)\lambda^m \tag{18}$$

Month:	Mean ratio(%):	Lower bound(%):	Upper bound(%):
January	22.64	22.29	22.99
February	26.24	25.80	26.68
March	28.22	27.71	28.72
April	31.95	31.28	32.62
May	32.69	31.98	33.39
June	32.32	31.66	32.98
July	33.43	32.72	34.13
August	31.72	31.06	32.37
September	29.45	28.92	29.97
October	24.03	23.63	24.43
November	22.71	22.36	23.06
December	22.67	22.32	23.02

Table 7: The ratio of average actual power generated and total wind power with lower and upper bounds of a 95% confidence interval

The $EP(V)$ was calculated by taking the mean of the power of the Weibull-simulated values for each month. The ratio was then calculated according to Equation 16. The confidence interval was, as previously, determined by the standard deviation and the p-value for the normal distribution of a 95% confidence interval according to Equation 7 and the standard deviation was calculated according to Equation 19.

$$\sigma = \frac{\sigma_{EP(V)}}{\sqrt{N} * P_{\text{tot}}(v)} \tag{19}$$

The ratios can be observed in Table 7. It can be noted that there is a moderate variability between the months in the ratio, where the summer months typically have a larger ratio.

3.7 g) The capacity factor and availability factor

In this part, the capacity factor and availability factor were calculated. The capacity factor is the ratio of the actual output over a time-period and the maximum possible output during the time-period, which is 9.5 MW per time-period for the Vestas V164. Typically, this is between 20% and 40% for wind turbines. The availability factor is the amount of time electricity is produced during a given time-period divided by the length of the period. This is typically above 90% for wind turbines. This ratio is very similar to the probability that was calculated in section 3.5, but it spans over a larger time-period.

In this task, the time-period was chosen to be 12 months. For the capacity factor, the vectors of the mean from the previous tasks were leveraged. The elements of the vectors were summed together and then divided by the time-period multiplied with the maximal output. This was done for the mean power output of the standard Monte Carlo and the truncated Monte Carlo, the mean using wind as a control variate, the mean using importance sampling, and the mean using antithetic sampling. The results can be observed in Table 8.

For the availability factor, the vector that stored the probability values from part 3.5 was utilized. In order to determine the amount of time power was produced during the whole year these values were summed together. This was then divided by the time-period. The result can be viewed in Table 9.

The value of 85% for the availability factor is quite reasonable, but on the low side as the value is typically above 90% for a wind turbine. This indicates that the site may not be optimal for a wind turbine, as one would prefer a larger availability factor.

Monte Carlo estimation method:	Capacity factor (%):
Standard Monte Carlo (2a)	39.3
Truncated Monte Carlo (2a)	39.3
Control Variate Monte Carlo (2b)	39.3
Importance sampling Monte Carlo (2c)	39.3
Antithetic sampling Monte Carlo (2d)	39.4

Table 8: Annual capacity factor of the different Monte Carlo estimation methods

Availability factor:	85.0%

Table 9: Annual availability factor of the wind turbine

The capacity factor varies considerably between the estimation methods, but they are within the typical interval. The average over all the estimation methods is 39.3% which is very high in the typical range.

Bearing this in mind, one can conclude that this site is not optimal for wind turbines, but it can however be satisfactory if there are no better options, as its capacity factor is very good and the availability factor is close to acceptable.

4 Part 3: Combined power production of two wind turbines

In this section of the assignment, an additional identical wind turbine was introduced to explore how the turbines affect each other.

4.1 a) The expected amount of combined power generated by both turbines

In this part, the problem was extended to determining the expected amount of combined power generated by both turbines together, E(P(V1) + P(V2)). In this case, it is possible to exploit the fact that they have the same distribution and thus have the same expected value. This implies that we can rewrite the problem as 2*E(P(V)) and that the problem is hence simplified to a one-dimensional plane.

This means that it is enough to estimate 2*E(P(V)) using importance sampling to solve this problem. It is possible to leverage the results from part 3.3 when determining this value. Since the parameter values for the Weibull distribution had been shifted from monthly values to yearly values, the correct probability density function was calculated. The new values were $k = 1.96$ and $\lambda = 9.13$ Then the power function was adjusted to represent two wind turbines and the importance sampling function was doubled as well to account for this change. These were then plotted together to ensure that they resembled each other, which can be observed in Figure 3. The figure shows that the distributions are quite similar, although they do not follow each other perfectly. Further, the ratio between them was plotted in Figure 4. Optimally, the ratio should be constant. In this case, the ratio fluctuated moderately, but the interval was still very small, between about 1.6 and 0.2. Thus, the instrumental distribution was deemed satisfactory.

Further, the confidence interval was calculated identically to in part 3.3, with the adjustment that $\phi(x) = 2 * P(x)$. The expected value and confidence interval can be observed in Table 10.

Expected value (W):	Lower bound (W):	Upper bound (W):	Width:
7134365	7084823	7183907	99 084

Table 10: The expected power generated by two wind turbines estimated with importance sampling

4.2 b) The covariance between the produced power in two identical wind power plants

The objective in this part of the assignment was to determine the covariance C(P(V1), P(V2)) between the produced power in two identical wind turbines receiving dependent winds. In order to calculate

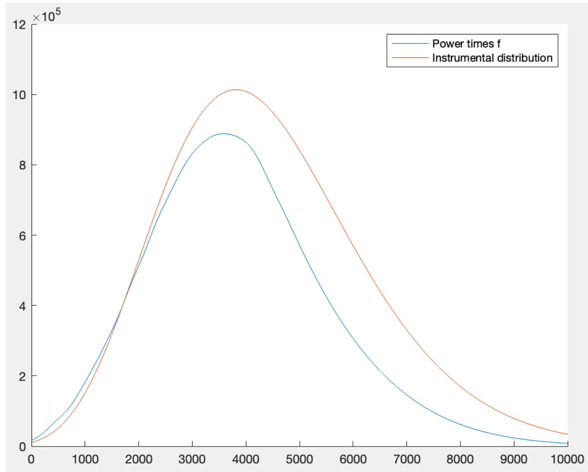

Figure 3: The instrumental distribution and $f(x) * \phi(x)$ in the bivariate case plotted against each other for January

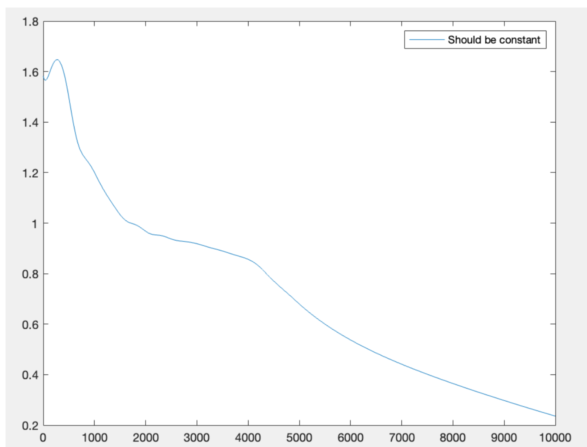

Figure 4: Plot of the ratio described by Equation 12 in the bivariate case for January

the covariance it had to be rewritten according to Equation 20 first.

$$C(P(V1), P(V2)) = E(P(V1)P(V2)) - E(P(V1))E(P(V2)) \tag{20}$$

When utilizing importance sampling the first term on the right-hand term can be rewritten according to Equation 21.

$$E(P(V1)P(V2)) = E_g(P(V1)P(V2)f(V1, V2)/g(V1, V2)) \tag{21}$$

In order to determine the covariance, the values of E(P(V1)) and E(P(V2)) had to be calculated. As they are identically distributed, the expected values are the same. This was determined through simulation in the same way as part 3.3 but with the appropriate values of λ and k. Additionally, the variance was determined which will be useful for part 4.3.

The next step was to determine the instrumental distribution to be able to perform importance sampling. In a 3D plot the bivariate Weibull distribution multiplied with P(V1)*P(V2) was plotted against different instrumental distributions, see Figure 5. It was difficult to find an instrumental distribution that followed the other curve well in 3D but a reasonable function was determined, see Matlab Equation 22. In the equation, V_1 and V_2 are the independent winds that are simulated with 100 samples of random Weibull distributions with the assumed values of k and λ-parameters. Then the mean was calculated using Equation 12, with P(V1)*P(V2) as ϕ, the bivariate Weibull distribution as f and the instrumental distribution as g.

$$g(V_1, V_2) = \text{gampdf}(V_1, 10, 1.3)' * \text{gampdf}(V_2, 25, 0.9) \tag{22}$$

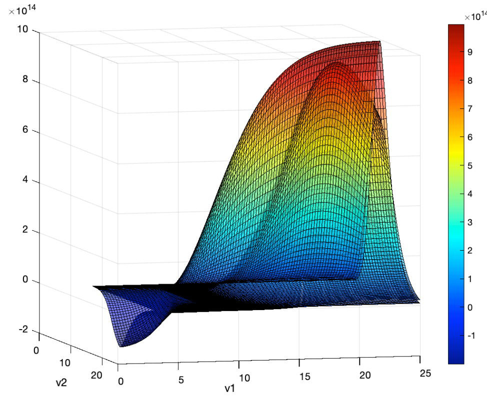

Figure 5: The instrumental distribution against bivariate Weibull distribution times the power of V1 times the power of V2

The covariance between the produced power from two identical wind power plants receiving dependent winds was estimated to approximately $9.14 * 10^{12}$ using importance sampling.

4.3 c) The variability in the amount of combined power

The purpose of this part was to estimate the variability V(P(V1)+P(V2)) in the amount of combined power generated by both wind turbines, as well as the standard deviation of the same case.

Here, it is possible to utilize the relationship between the variances and covariance described in Equation 23. Further, it can be noted that the relationship between the variance and standard deviation is described by Equation 24.

$$V(P(V1) + P(V2)) = V(P(V1)) + V(P(V2)) + 2 * C(P(V1), P(V2)) \tag{23}$$

$$D(P(V1) + P(V2)) = \sqrt{V(P(V1) + P(V2))} \tag{24}$$

The covariance had already been determined in Part 4.2, so this information could be leveraged in Equation 23. What remained was then to estimate $V(P(V1))$ as it is known that the two turbines are identical, so the variances will be equivalent. This practice has been done in multiple different variance-reducing ways in this assignment for other λ and k-values, but here it was determined sufficient to only use standard Monte Carlo to estimate the variance since it is the simplest. Therefore, 100,000 samples were created with the correct λ and k-values. The variance among them was then used for this estimate as V(P(V1)) and V(P(V2)).

The obtained values were then inserted into Equation 23 and 24, respectively. The resulting values can be observed in Table 11.

Variance:	42 460 665 368 462
Standard deviation:	6 516 184W

Table 11: The variance and standard deviation of the combined power production using importance sampling

4.4 d) The probability that the combined power generated by the two turbines is greater or lesser than half of their installed capacity

In this part the probability that two power turbines in the same area produce more than 9.5MW power together and the probability that they produce less than 9.5MW were calculated. The first probability is

$$P(P(V_1) + P(V_2) > 9.5MW)$$

. Importance sampling was chosen as the variance-reducing technique in this part for both probabilities. To find a suitable instrumental distribution the probability was plotted in a 3D plot together with different possible instrumental distributions. After analyzing the graph and testing different distributions, it was decided to use two gamma functions multiplied with each other, both with shape parameter 10 and scale parameter 1.3, as the instrumental distribution, which can be observed in Equation 25. See Figure 6 for the resulting instrumental distribution together with the probability. It is displayed from a different angle in Figure 7.

$$g(V) = \text{gampdf}(V_1, 10, 1.3) * \text{gampdf}(V_2, 10, 1.3) \tag{25}$$

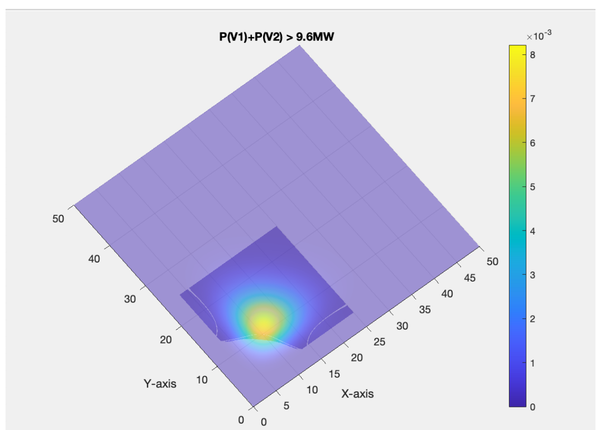

Figure 6: Instrumental distribution and probability for $P(P(V_1) + P(V_2) > 9.5MW)$

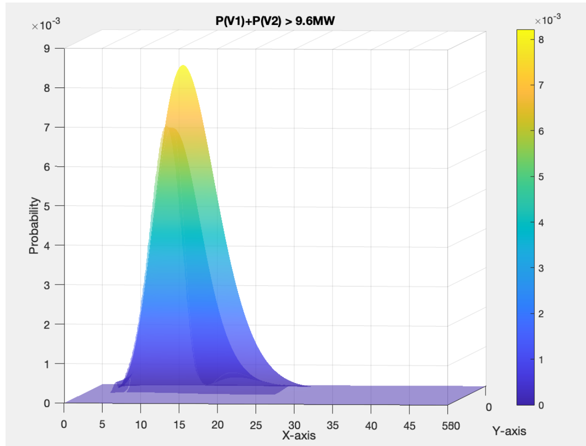

Figure 7: Instrumental distribution and probability from a different angle for $P(P(V_1) + P(V_2) > 9.5MW)$

The resulting probability was calculated with random gamma values as input, applying the importance sampling technique previously described generalized for the bivariate case. This was calculated 100 times to get a confidence interval of the probability, and the confidence interval was calculated using Equation 7. The resulting probability with its confidence interval was: 0.8132 +/- 0.1102.

For the opposite probability, the same procedure was performed, but with the probability

$$P(P(V_1) + P(V_2) < 9.5MW)$$

. A suitable instrumental distribution was found to be two Weibull distributions with identical parameters, given by the marginal distributions of V_1 and V_2, multiplied with each other, see Equation 26. See Figure 8 for the resulting instrumental distribution together with the probability, where the instrumental distribution is gridded. It can be observed from a different angle in Figure 9.

$$g(V) = \text{wblrnd}(V_1, 9.13, 1.96) * \text{wblrnd}(V_2, 9.13, 1.96) \tag{26}$$

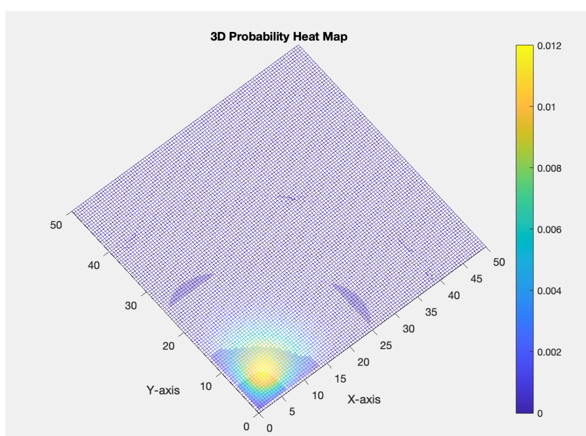

Figure 8: Instrumental distribution and probability for $P(P(V_1) + P(V_2) < 9.5MW)$

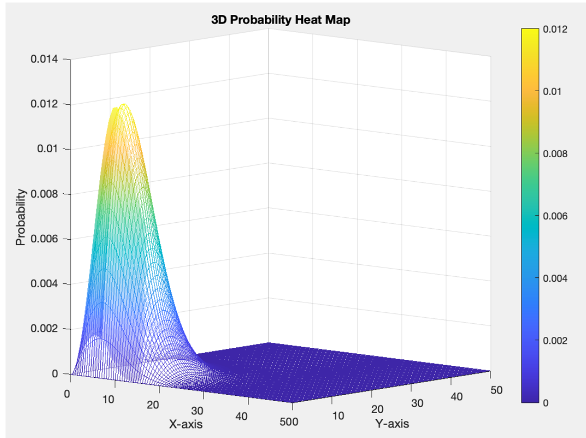

Figure 9: Instrumental distribution and probability from a different angle for $P(P(V_1) + P(V_2) < 9.5MW)$

The resulting confidence interval was calculated by running the simulation 100 times and calculating the confidence interval from the values of the simulations with Equation 7. This was done similarly to what was previously described for the first probability. The resulting probability with its confidence interval was: 0.1970 +/- 0.0024.

The probabilities sum up to around one in this case. However, since the case where the sum of the power of the turbines is exactly 9.5MW is not in any of the probabilities, the sum of the probabilities should be a bit less than one. Further, the values are simulated with 100 different samples and the average is then taken. This means that there can be small variations due to the randomness of the simulation.

5 Final words

In this assignment, various statistical techniques to estimate the expected power production of a wind turbine were explored, using Monte Carlo simulation, importance sampling, control variates, and antithetic sampling. By leveraging different variance reduction methods, we were able to enhance the precision of our estimates and compare the effectiveness of these approaches.

This assignment highlights the importance of advanced statistical methods in fields like renewable energy modeling, demonstrating how computational techniques can improve decision-making.

All MATLAB implementations, including the main script (proj1.m) and supporting files, have been submitted via email, and this report has been uploaded in PDF format to CANVAS.