NTNU

Group Assignment

Group 2023
Members:

Payam	Rajabzadeh Kanafi
Fahmid	Ul-Hasan
Pierre	Nussbaum

Course Number and Name TBA5150 Geohazards and Risks

Table of Contents

List of Figures

List of Tables

I Statistical analysis

The first step to solve this problem is to define it. We are given the following expression for the factor of safety:

$$F = \frac{\epsilon S_u \alpha}{\frac{1}{4}\gamma R(\alpha - sin\alpha)sin\alpha + \eta q sin\alpha} \tag{1}$$

We now need to determine the distribution of those parameters. Some are given but we need to analyse the data given to find the distribution of γ and S_u.

Thus, for each zone, we have plotted the q-q plots between the data given and several distribution (cf. A). We compared the distribution by calculating the correlation of each graph. The following table shows the best distributions found for each zone:

Param.	**Zone 1**	**Zone 2**	**Zone 3**	**Zone 4**	**Zone 5**	**Zone 6**	**Zone 7**
S_u	N/Log	N/Log/G	N/Log	N/Log/G	N/Log/G	N/Log	N/Log
γ	N/U	N/U	N	N	N/U	G	N

Table 1: Best distribution for S_u and γ for each zone
N=Normal, Log=Lognormal, G=Gamma, U=Uniform

Even though we can wonder if the statistical approach to γ is relevant with so little data, we can see that the normal distribution fits the best to the data. For S_u, we considered choosing lognormal distribution but the probability of failure calculated in the next questions were much lower compared to a normal distribution. Thus, we assumed a normal distribution for S_u as a conservative approach. The statistical features of all the parameters are gathered in table 2.

Rand. Var.	**Dist.**	**Par.**	**Zone 1**	**Zone 2**	**Zone 3**	**Zone 4**	**Zone 5**	**Zone 6**	**Zone 7**
ϵ	N	μ_ϵ	1	1	1	1	1	1	1
		σ_ϵ	0.1	0.1	0.1	0.1	0.1	0.1	0.1
S_u (kPa)	N	μ_{S_u}	46.86	56.24	44.60	51.55	37.07	49.83	54.00
		σ_{S_u}	8.47	7.56	9.73	5.19	6.06	9.15	11.09
α (°)	N	μ_α	60	60	60	60	60	60	60
		σ_α	9	9	9	9	9	9	9
γ	N	μ_γ	18.40	18.34	17.30	18.78	18.72	19.64	16.88
		σ_γ	1.02	2.04	2.73	1.98	1.40	2.17	2.59
R	N	μ_R	8	8	8	8	8	8	8
		σ_R	1.6	1.6	1.6	1.6	1.6	1.6	1.6
η	N	μ_η	0.1	0.1	0.1	0.1	0.1	0.1	0.1
		σ_η	0.01	0.01	0.01	0.01	0.01	0.01	0.01
q (kPa)	N	μ_q	10	10	10	10	10	10	10
		σ_q	2	2	2	2	2	2	2

Table 2: Random parameter distribution for every zones

II FOSM

We assumed all parameters are independent and we used the FOSM approximation to compute the mean and standard deviation of the factor of safety.

	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7
mean	7.3899	8.8948	7.4192	7.9862	5.7598	7.42388	9.17729
std	3.1090	3.6597	3.3659	3.2077	2.3997	3.2149	4.0863

Table 3: Statistical features of the factor of safety F

We considered a performance function such as :

$$g = F - 1 \tag{2}$$

Thus $\mu_g = \mu_F - 1$ and $\sigma_g = \sigma_F$. The reliability index of g (β) can be deduced by dividing the mean by the standard deviation. Then, the probability of failure is calculated using ϕ function from the package scipy.stats: $p_f = \phi(-\beta)$.

	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7
Mean	6.3899	7.8948	6.4192	6.9862	4.7598	6.42388	8.17729
Standard deviation	3.1090	3.6597	3.3659	3.2077	2.3997	3.2149	4.0863
Reliability index	2.0553	2.1573	1.9071	2.1779	1.9835	1.9981	2.0012
Failure probability	0.019924	0.015493	0.028251	0.014706	0.023655	0.022852	0.02269

Table 4: Statistical features of the performance function g and probability of failure using FOSM method

We now consider another performance to improve the degree of linearity of the problem:

$$g = R - S = \epsilon S_u \alpha - (0.25\gamma R(\alpha - sin\alpha)sin\alpha + \eta q sin\alpha)$$

Now, $\mu_g = \mu_R - \mu_S$ and $\sigma_g = \sigma_R - \sigma_S$. By running the same code as before with the new performance function, we have the following results:

	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7
Mean	42.429	52.272	40.407	47.220	32.083	45.153	50.3959
Standard deviation	11.059	11.490	12.012	9.2085	8.0028	11.897	14.156
Reliability index	3.8365	4.5492	3.3640	5.1280	4.009	3.7954	3.5600
Failure probability	6.240e-5	2.693e-6	3.841e-4	1.464e-7	3.049e-5	7.372e-5	1.854e-4

Table 5: Statistical features of the new performance function g and probability of failure using FOSM method

Various performance functions yield distinct probabilities of failure. This discrepancy arises from FOSM's reliance on a linear approximation of the performance function solely at its mean to calculate the reliability index. Consequently, if the linear approximation at the mean fails to accurately represent the performance function, FOSM may either overstate or underestimate the reliability index.

III FORM

The FORM method has been done using Openturns package setting absolute error value to 0.001. The event considered by Openturns is F>1, which is equivalent to using the performance function defined previously (cf. equation 2).

	Zone 1	**Zone 2**	**Zone 3**	**Zone 4**	**Zone 5**	**Zone 6**	**Zone 7**
Failure probability	3.35e-6	2.19e-9	7.23e-5	2.27e-10	5.68e-6	5.03e-6	1.28e-6

Table 6: Failure probability for every zones using FORM method

The sensitivity coefficients for random parameters are presented in B. We can see that the main parameter for every zone, except zone 4, is the shear strength.

The following table shows the coordinates of the design point using the FORM method.

	Zone 1	**Zone 2**	**Zone 3**	**Zone 4**	**Zone 5**	**Zone 6**	**Zone 7**
ϵ	0.937	0.842	0.962	0.710	0.897	0.938	0.964
α	1.278	1.460	1.202	1.582	1.370	1.278	1.193
R	9.476	10.977	8.945	12.198	10.203	9.476	8.893
q	10.173	10.195	10.123	10.205	10.168	10.137	10.124
η	0.100	0.100	0.100	0.101	0.100	0.100	0.100
S_u	12.178	22.441	9.881	36.183	16.352	13.497	9.375
γ	18.708	20.889	18.646	22.306	19.585	20.897	18.063

Table 7: Coordinates of the design point for each zone

The sensitivity of the parameter on the failure probability is presented in appendix B. We can see that the two most sensitive factors are S_u and α.

IV Monte Carlo

The graphs related to the Monte Carlo simulations are presented in appendix C The following table shows the results :

	Ns	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7
Mean	10^6	8.26	10.0	8.42	9.0	6.45	8.40	10.38
Standard deviation	10^5	3.82	4.47	4.12	3.92	2.94	3.96	4.96
Probability of failure	10^8	4.22e-6	0	7.84e-5	0	6.28e-6	5.78e-6	1.34e-5
CoV of the estimate pf	10^8	0.050	-	0.011	-	0.040	0.042	0.27

Table 8: Results of the Monte Carlo simulations

The MC for the probability of failure has been done ith 10^8 simulations but it is still not enough to get a failure point for zones 2 and 4 (see figure C.4). This is a limitation of the MC simulation for rare event. As we have seen in problem 3, the failure probabilities for these zones with the FORM approach are 2.10^{-9} and 2.10^{-10}. Given this, it requires respectively 5.10^{10} and 5.10^{11} for a coefficient of variation of 0.1 which is unreasonably high for our computers. The probability of failure of zones 2 and 4 are then set to 0.

We can see that the different methods give very different results and in that case over-estimate or underestimate the failure probability. This is due to the non-linearity of the system. FORM method is more accurate but also harder to realize in practice. The assumptions of this method might not be conservative since it can underestimate the risk compared with the MC approach. That can be an issue for some application. Yet, compared to MC method, it doesn't have any limitation for rare event models.
Finally, MC approach is simple and doesn't require many assumptions but it's longer to compute and it is not a deterministic method so the result varies from on simulation to another. Moreover, MC method can be inconsistent for rare events as we can see for zones 2 and 4.

	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7
FOSM	0.019924	0.015493	0.028251	0.014706	0.023655	0.022852	0.0226
FOSM2	6.240e-5	2.693e-6	3.841e-4	1.464e-7	3.049e-5	7.372e-5	1.85394e-4
FORM	3.35e-6	2.19e-9	7.23e-5	2.27e-10	5.68e-6	5.03e-6	1.28e-6
MC	4.22e-6	0	7.84e-5	0	6.28e-6	5.78e-6	1.34e-5

Table 9: Probability of failure for FOSM, FORM and MC methods

To overcome the drawbacks of the MC method, we adapt the method for rare event. The idea is to shift the mean of the parameters to over estimate the failure probability in order to have more failure points detected. The proposal distribution is defined from the original distribution where the mean of S_u, ϵ and α is reduced by one standard deviation while the mean of γ, R, η and q was increase by one standard deviation. The probability of failure of the proposal distribution is then weighted by a factor to get the the probability of failure of the original distribution.

This method gives the following probability:

	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7
MC sampling	3.66e-6	3.06e-10	8.96e-5	4.55e-11	6.73e-6	6.26e-6	1.37e-5

Table 10: Results of the MC for rare event

Yet, even with this method, the result varies from one simulation to another and it is hard to validate the code so the results of this simulation will not be used in next sections.

More information in Appendix C.

Validation of the code

All the code (FORM, FOSM and MC) have been validated using a simple example:

$$F = \epsilon + S_u + \alpha + \gamma + R + \eta + q - 85$$

. The result is exactly the same with every method.

	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7
Mean	0.404	9.726	-2.956	5.474	-9.059	4.618	6.037
Standard deviation	8.908	8.239	10.43	6.119	6.727	9.751	11.68
Reliability index	0.045	1.181	-0.283	0.895	-1.347	0.474	0.517
Failure probability	0.482	0.119	0.612	0.186	0.911	0.318	0.303

Table 11: Validation of the code

V System reliability

The problem can be represented by the following scheme:

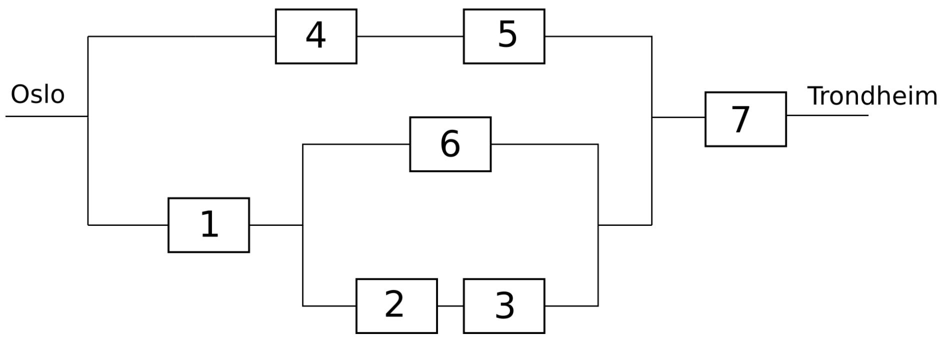

Figure 1: Represention of the problem

The minimal cut of this system are :

$$[7], [1,4], [1,5], [2,4,6], [2,5,6], [3,4,6], [3,5,6]$$

The system is then equivalent to the following one:

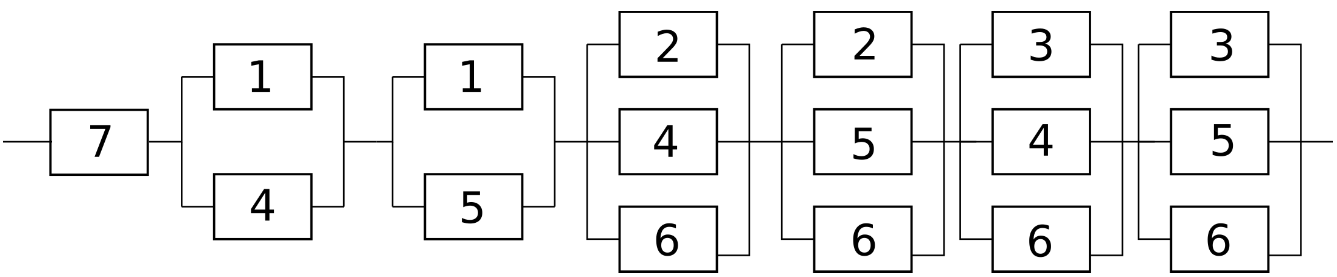

Figure 2: Equivalent system

The probability of failure of the system is then equal the the product of the failure probability of all subsystems.

$$\begin{aligned} \text{Pf} = \quad & 1 - (1 - Pf_7)(1 - Pf_1 Pf_4)(1 - Pf_1 Pf_5)(1 - Pf_2 Pf_4 Pf_6) \\ & (1 - Pf_2 Pf_5 Pf_6)(1 - Pf_3 Pf_4 Pf_6)(1 - Pf_3 Pf_5 Pf_6) \end{aligned}$$

We can also consider an event tree to describe the situation (cf. Appendix D). The failure probability is then 1 minus the sum of the probabability of succesful outcome:

$$\begin{aligned} \text{Pf} = \quad & 1 - [(1 - Pf_7)(1 - Pf_1)(1 - Pf_6) + (1 - Pf_7)(1 - Pf_1)Pf_6(1 - Pf_2)(1 - Pf_3) + \\ & (1 - Pf_7)(1 - Pf_1)Pf_6(1 - Pf_2)Pf_3(1 - Pf_4)(1 - Pf_5) + \\ & (1 - Pf_7)(1 - Pf_1)Pf_6 Pf_2(1 - Pf_4)(1 - Pf_5) + \\ & (1 - Pf_7)Pf_1(1 - Pf_4)(1 - Pf_5)] \end{aligned}$$

Both method gives the same results presented in table 12. FOSM method gives a higher value for the failure probability of the system, which is due to the high value of failure probability for each zone. FORM and MC methods give again similar results.

Given the costs detailed in the next section, we can estimate the expected cost if Yutao choose to take the car:

$$E = 1200 * (1 - p_{f_S}) + 46200 * p_{f_S}$$

where p_{f_S} if the probability of failure of the system presented in table 12.

	FOSM	**FOSM2**	**FORM**	**MC**
Failure probability of the system	$2.35\ 10^{-2}$	$1.85\ 10^{-4}$	$1.28\ 10^{-5}$	$1.34\ 10^{-5}$
Expected cost (kr)	2257.5	1208.33	1200.58	1200.60

Table 12: System failure probabilities (p_{f_S}) and expected cost for every method

Yet, this is not the right model of the problem since Yutao has to take the plane if there is a landslide on the chosen route, he can't turn back and take another route. Thus, the expected cost is in fact underestimate. Moreover, it doesn't estimate the probability of failure of each road so we can't conclude which route to take. A new approach is needed to solve the problem.

VI Conclusion

Yutao has to choose between 4 options: take the plane right away or take his car and try to overcome the danger of the road by choosing one of the 3 routes. Each possibility has a cost and we now need to find the best solution with the lowest expectation of cost.

This problem has 3 outcomes:

- Taking the plane right away has a total cost of 26200 Kr: 20000 kr for the plane, 5000 kr for the parking and traveling back to the airport and 1200 kr to go from Oslo to Trondheim by car.
- The cost of traveling by car from Oslo to Trondheim without any problem is 1200 kr.
- Finally, the cost if Yutao is stopped by a landslide is 46200 kr: 40000 kr for the plane, 5000 kr for the parking and the travel to take the car back and 1200 for the gas for the travel between Oslo and Trondheim.

If p_{f_R} is the probability of failure of one route with one method, the expected expense can be calculated:

$$E = 1200 * (1 - p_{f_R}) + 46200 * p_{f_R}$$

First, we need to calculate the probability of failure of each route. Each route can be represented as a serial system of its danger zones. Then the probability of failure of each route is:

$$p_{f_R} = 1 - \prod (1 - p_{f_Z})$$

p_{f_Z} is the probability of failure of each danger zone of the route.

	FOSM	**FOSM2**	**FORM**	**MC**
Route 1	0.084	6.4 10^{-4}	8.9 10^{-5}	9.6 10^{-5}
Route 2	0.064	3.2 10^{-4}	2.1 10^{-5}	2.3 10^{-5}
Route 3	0.060	2.2 10^{-4}	1.9 10^{-5}	2.0 10^{-5}

Table 13: Probability of failure for each route

The following table shows the expected cost of each possibility for every method.

	FOSM	**FOSM2**	**FORM**	**MC**
Plane	26200	26200	26200	26200
Route 1	4963.92	1228.55	1203.98	1204.32
Route 2	4082.29	1214.47	1200.96	1201.05
Route 3	3893.86	1209.72	1200.83	1200.89

Table 14: Expectation of cost in kr

Conclusion The comprehensive evaluation of various methodologies consistently indicates that Route 3 emerges as the optimal solution, demonstrating the lowest anticipated cost. Despite the thorough analysis, it is imperative to acknowledge the inherent uncertainty in decision-making processes. The ultimate determination rests with the decision-maker, who possesses the discretion to assess the risks and benefits associated with adopting Route 3 as the preferred course of action.

Appendix

A Distribution of S_u and γ

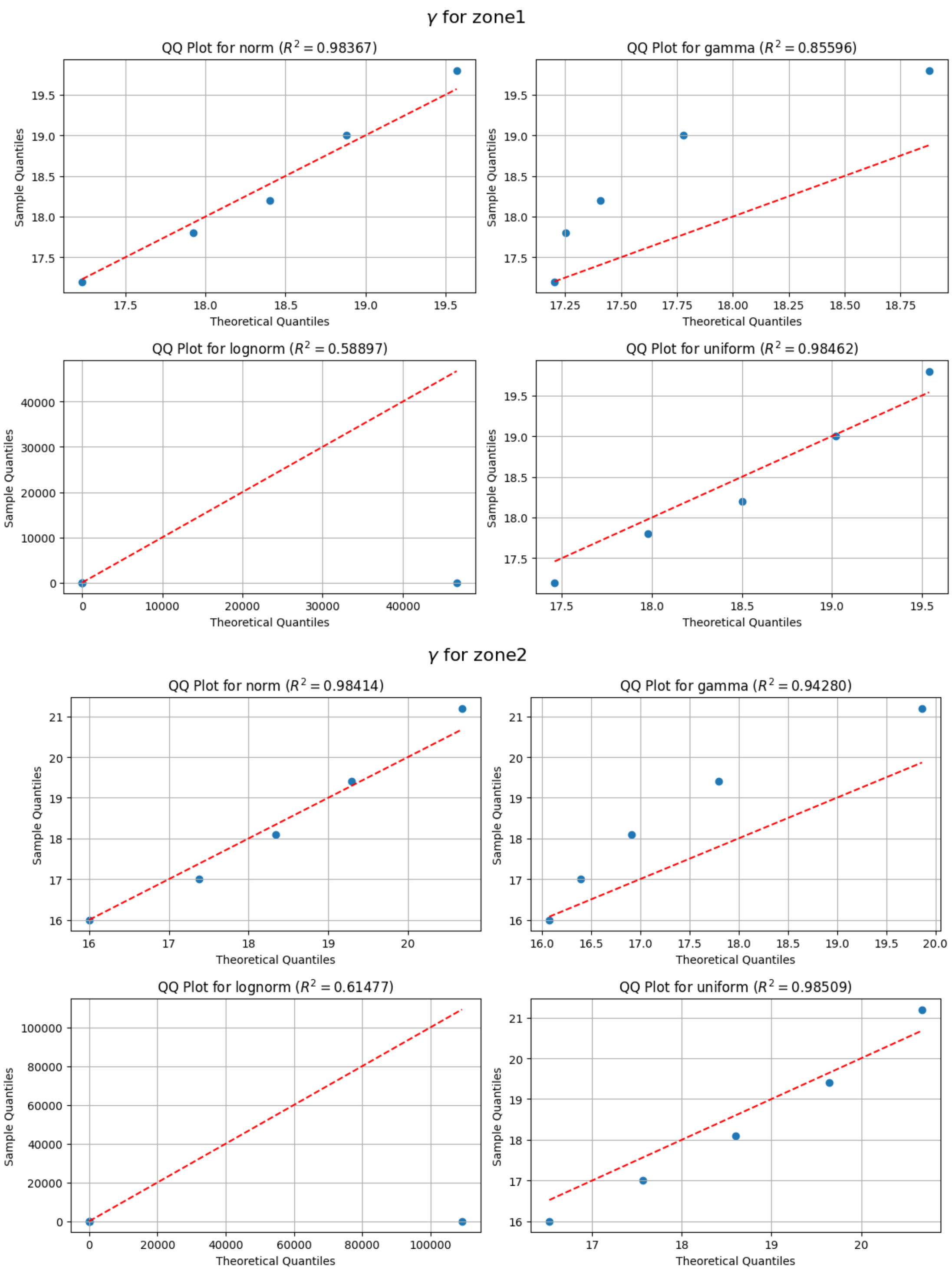

Figure A.1: Distrbutions for γ for zones 1 and 2

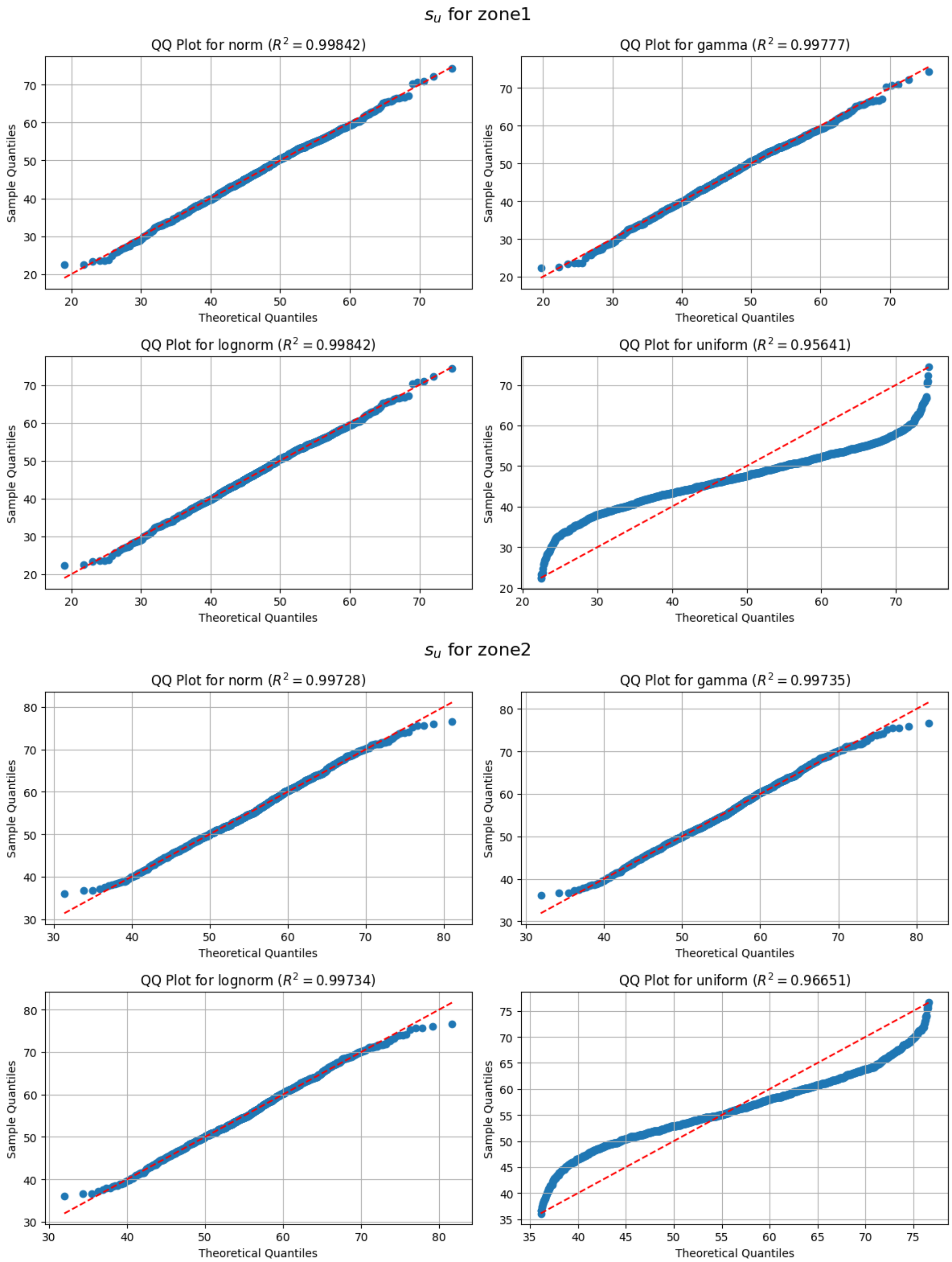

Figure A.2: Distrbutions for S_u for zones 1 and 2

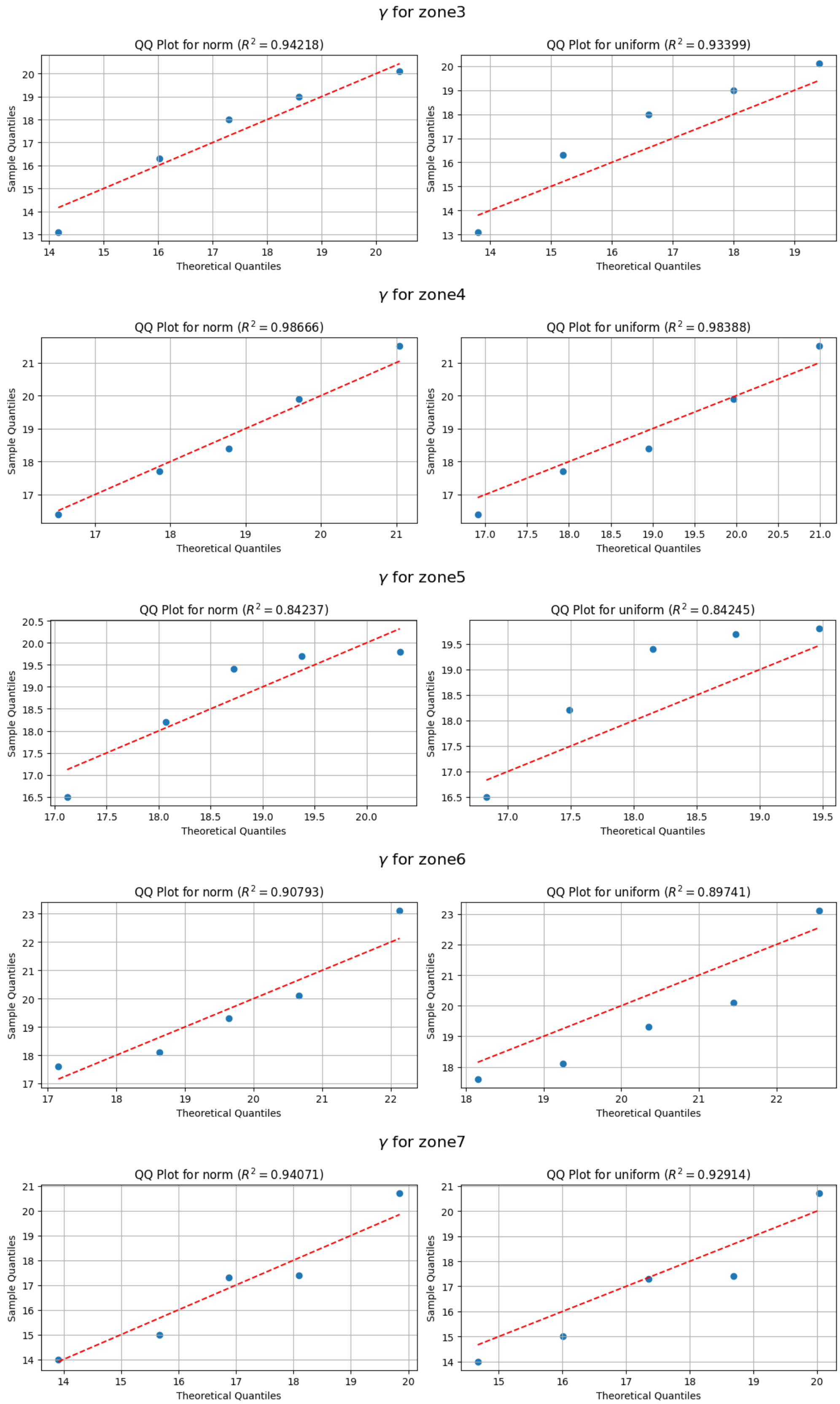

Figure A.3: Distribution best suited for gamma for the rest of the zones

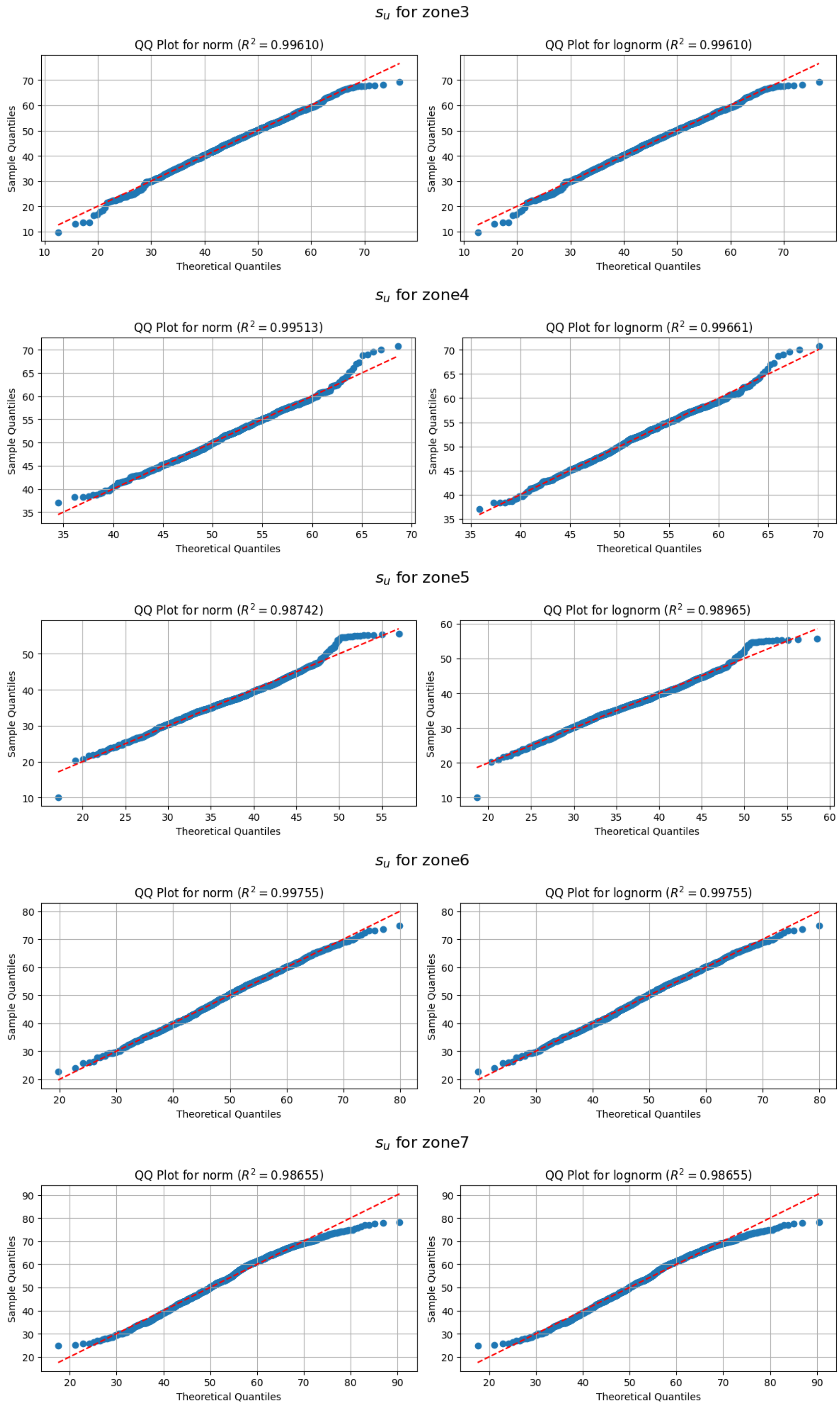

Figure A.4: Distribution best suited for S_u for the rest of the zones

B Sensitivity of the parameters

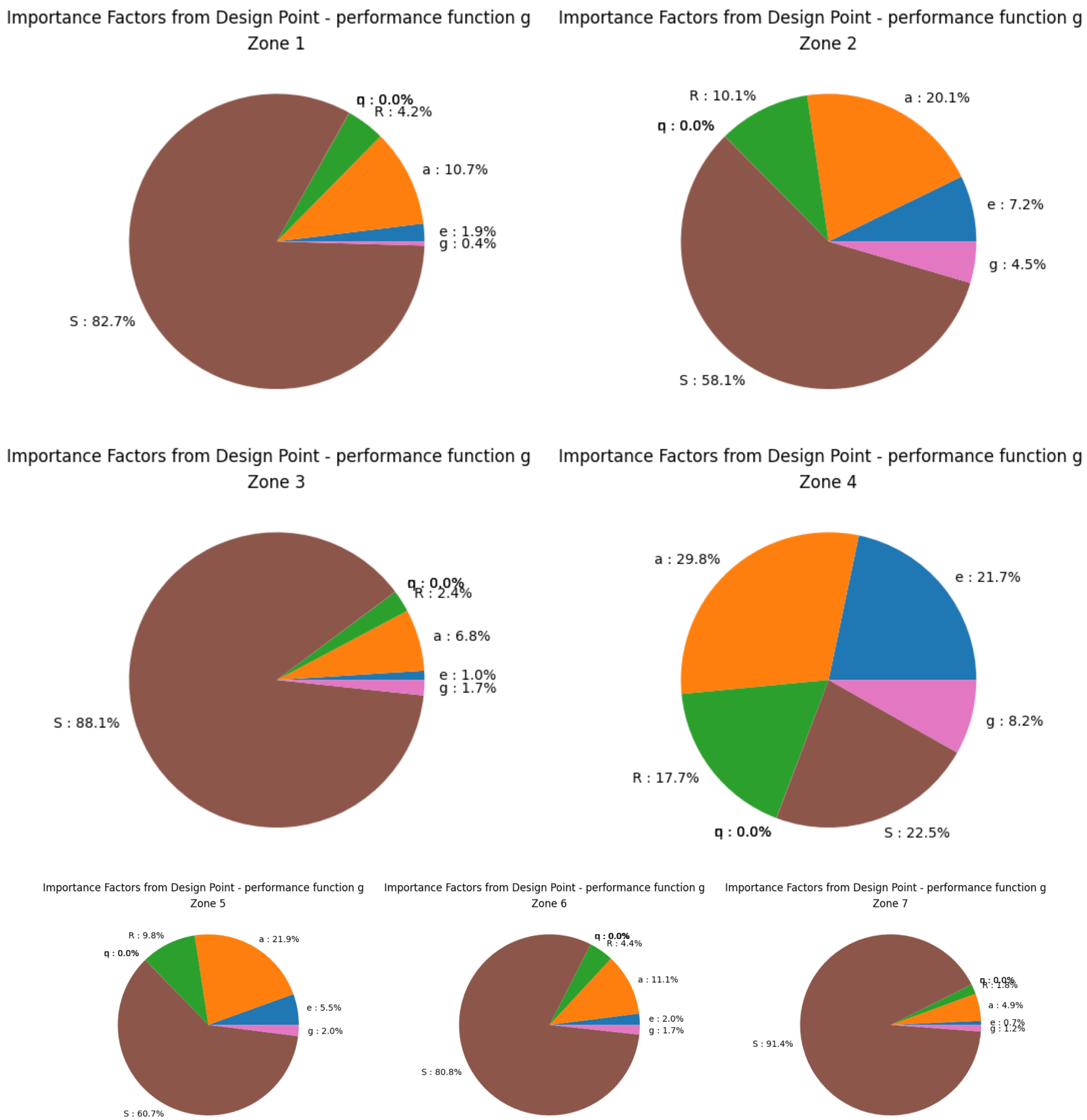

Figure B.1: Importance factors from design point of the FORM method

	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7
ϵ	1.9	7.2	1.0	21.7	5.5	2.0	0.7
S_u	82.7	58.0	88.1	22.5	60.7	80.8	91.4
α	10.7	20.0	6.8	29.8	21.9	11.1	4.9
γ	0.4	4.5	1.7	8.1	2.0	1.7	1.2
R	4.2	10.0	2.4	17.7	9.8	4.3	1.8
η	0.008	0.004	0.003	0.009	0.01	0.005	0.004
q	0.04	0.02	0.02	0.03	0.03	0.02	0.02

Table 15: Importance factor for each parameter in %

C MC simulation results

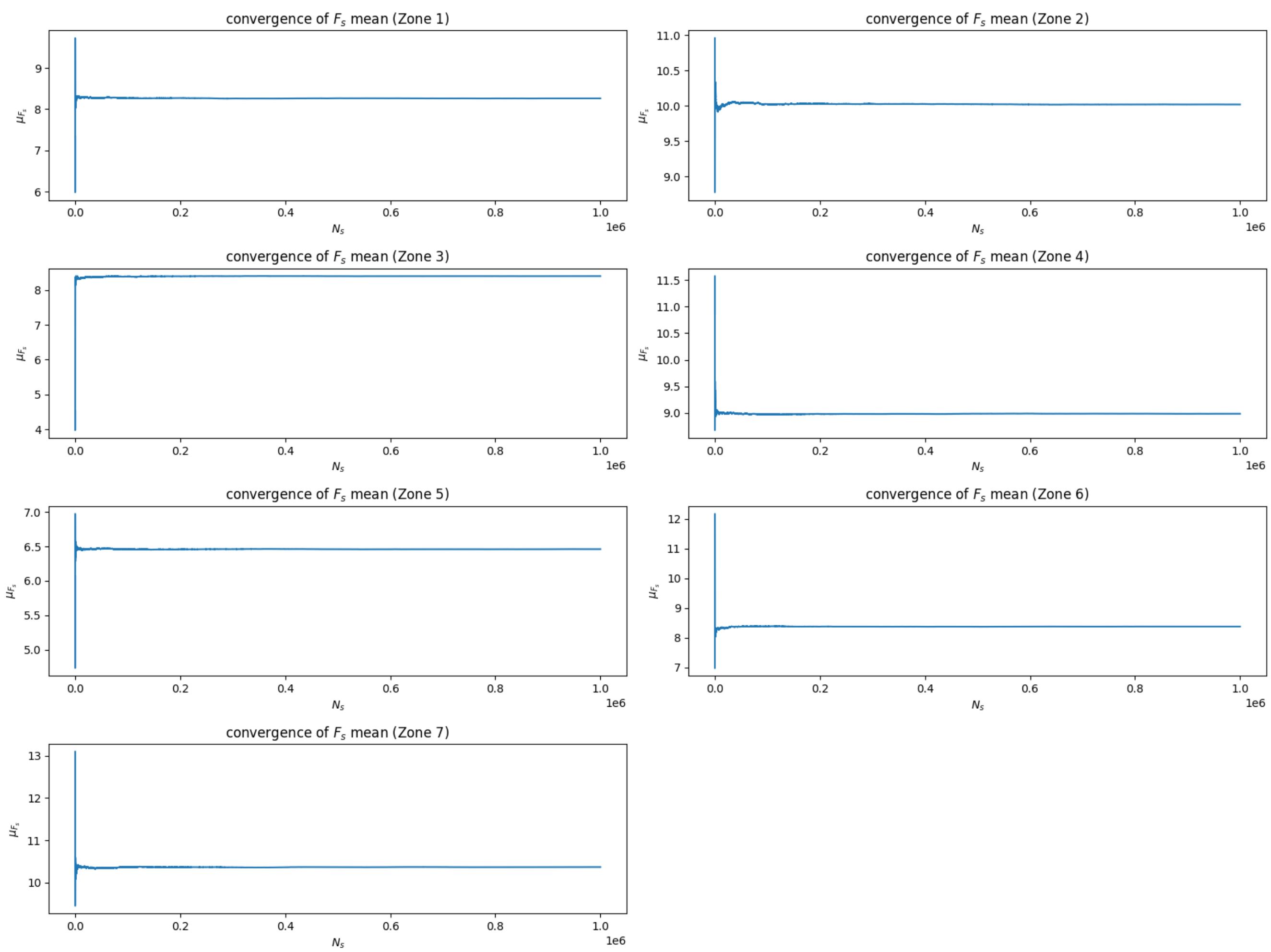

Figure C.1: Convergence of the mean of F

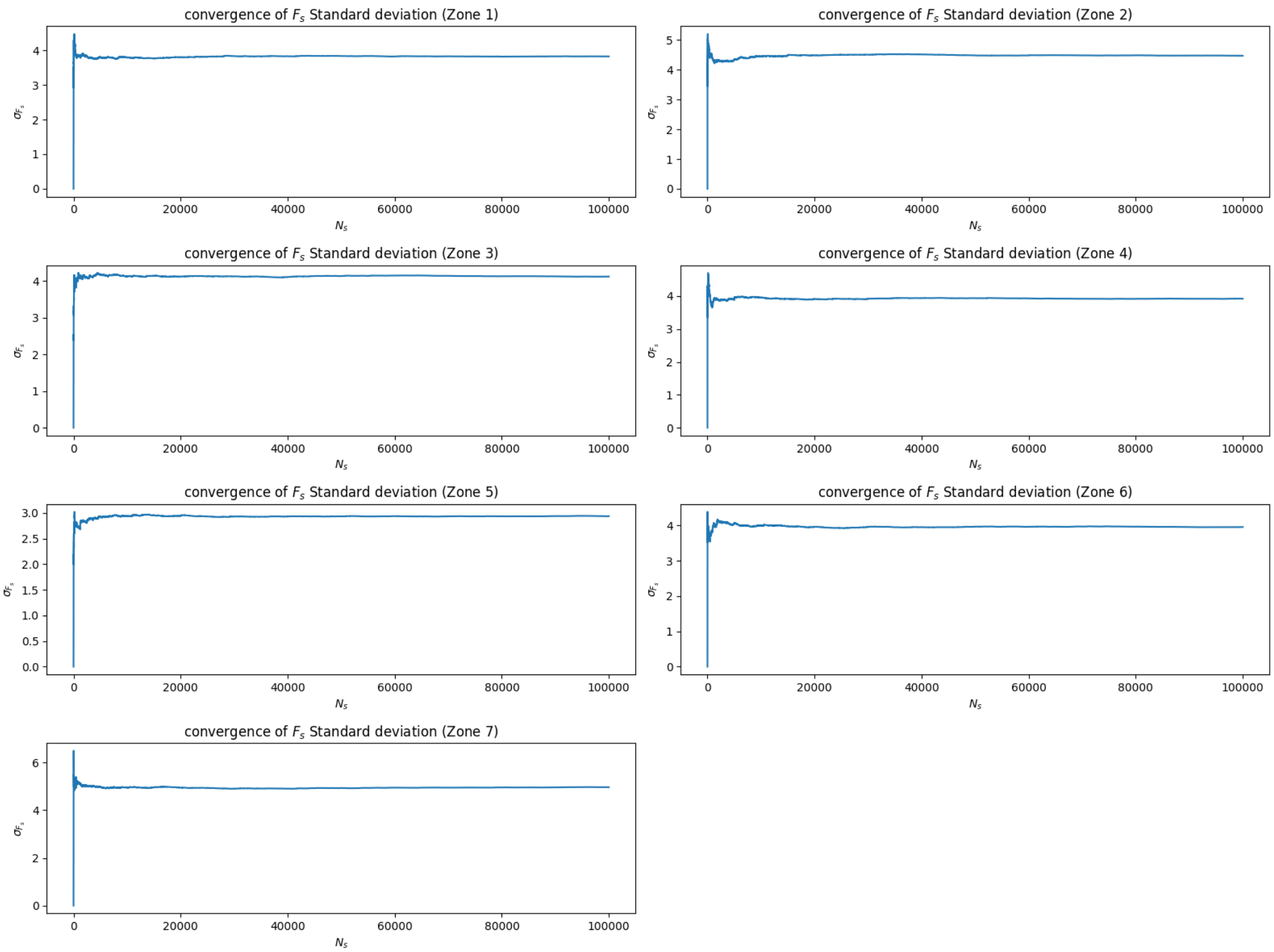

Figure C.2: Convergence of the standard deviation of F

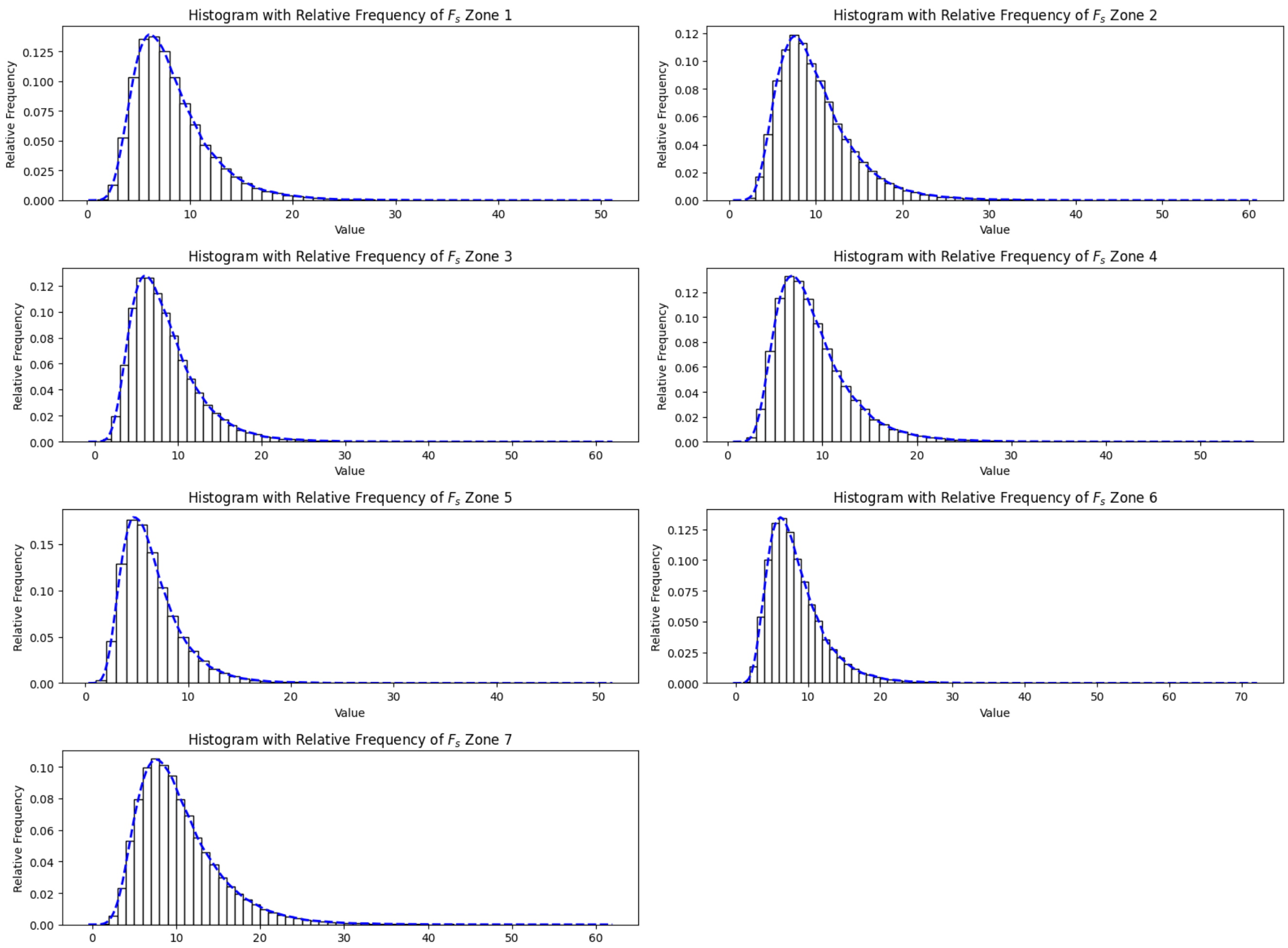

Figure C.3: Distribution of F according Monte Carlo simulation

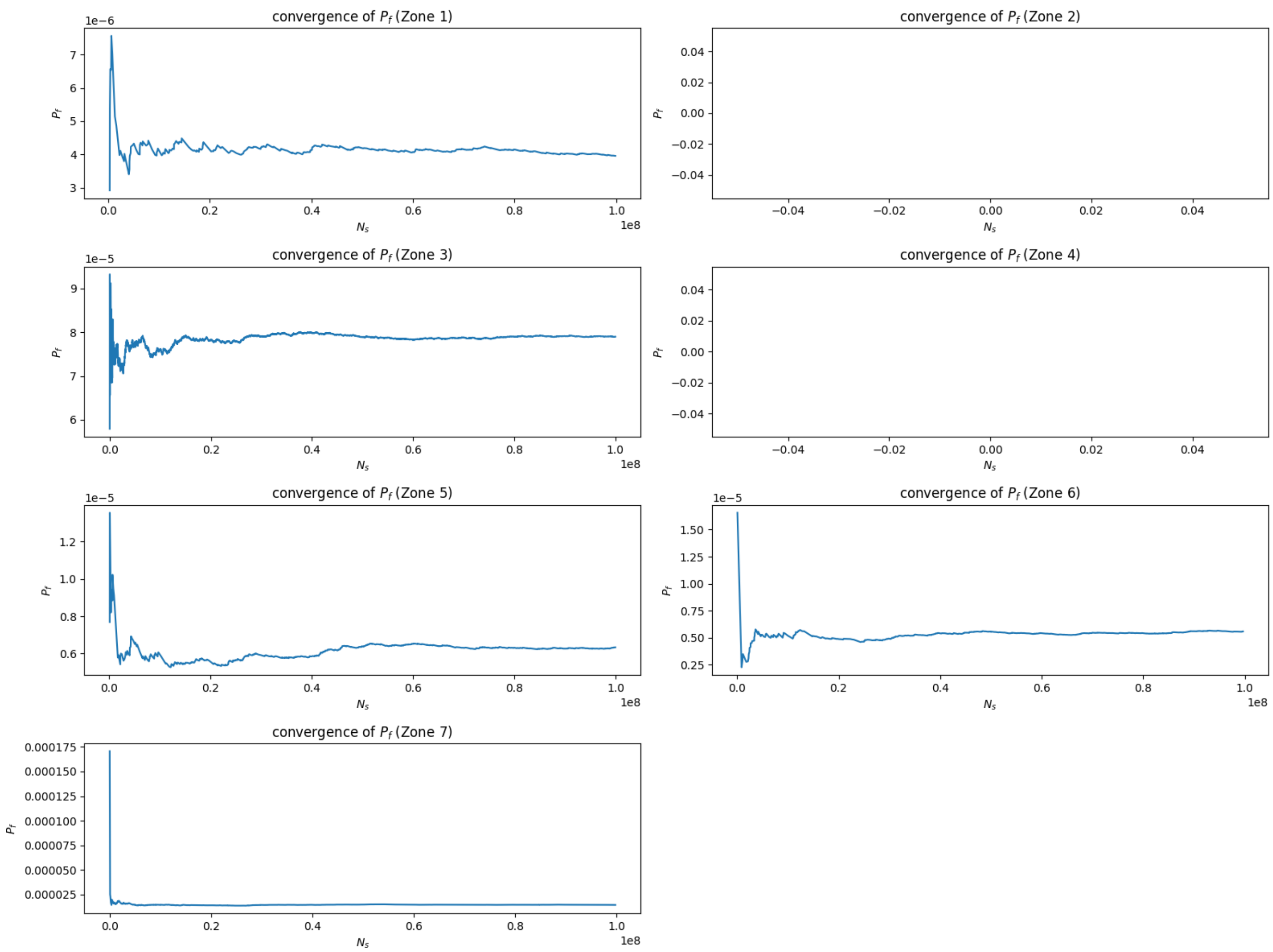

Figure C.4: Convergence of the probability of failure

Code of MC for rare event

```
for zone in zones:
  z1=mean_parameters.loc[:,zone]
  z2=std_parameters.loc[:,zone]


  # Define parameters for the original distribution
  means = [z1[0],z1[1],np.radians(z1[2]),z1[3],z1[4],z1[5],z1[6]]
  stds = [z2[0],z2[1],np.radians(z2[2]),z2[3],z2[4],z2[5],z2[6]]

  y=1
  propmeans= [z1[0]-y*z2[0],z1[1]-y*z2[1],np.radians(z1[2]-y*z2[2]),
              z1[3]+y*z2[3],z1[4]+y*z2[4],z1[5]+y*z2[5],z1[6]+y*z2[6]]

  # Define the function of interest
  def target_function(x):
      return ((x[0]*x[1]*x[2])/(1/4*x[3]*x[4]*(x[2]-np.sin(x[2]))
                              *np.sin(x[2])+x[5]*x[6]*np.sin(x[2])))-1
  # Define the original distribution (P)
  def original_distribution_sample(size):
      return np.random.normal(loc=means, scale=stds, size=(size, len(means)))

  # Define the proposal distribution (Q)
  def proposal_distribution_sample(size):
      return np.random.normal(loc=propmeans, scale=stds, size=(size, len(means)))

  # Number of samples
  num_samples = 10**7

  # Generate samples from the proposal distribution
  q_samples = proposal_distribution_sample(num_samples)

  # Calculate the importance weights
  weights = np.prod(norm.pdf(q_samples, means, stds) / norm.pdf(q_samples,
                              propmeans, stds), axis=1)
  is_less_than_zero = target_function(q_samples.T)<=0
  estimate = np.sum(weights * is_less_than_zero) / num_samples

  print(f"Estimated probability of failure for {zone} is {estimate}")
```

Limits of the MC for rare event
As explained in Section IV, we shift the mean of the parameter by y time the standard deviation (see lines 12 and 13 of the code above). Different values of y were tested but it seems that the accuracy is lower when y is higher than 1. Thus, a significant number of simulation is still needed to find failure point for zones 2 and 4 and it limits the convergence of the failure probability. Moreover, relatively important variation were noticed when running the same code several time in a row. Unfortunately, no solution was found for this problem.

D Event tree

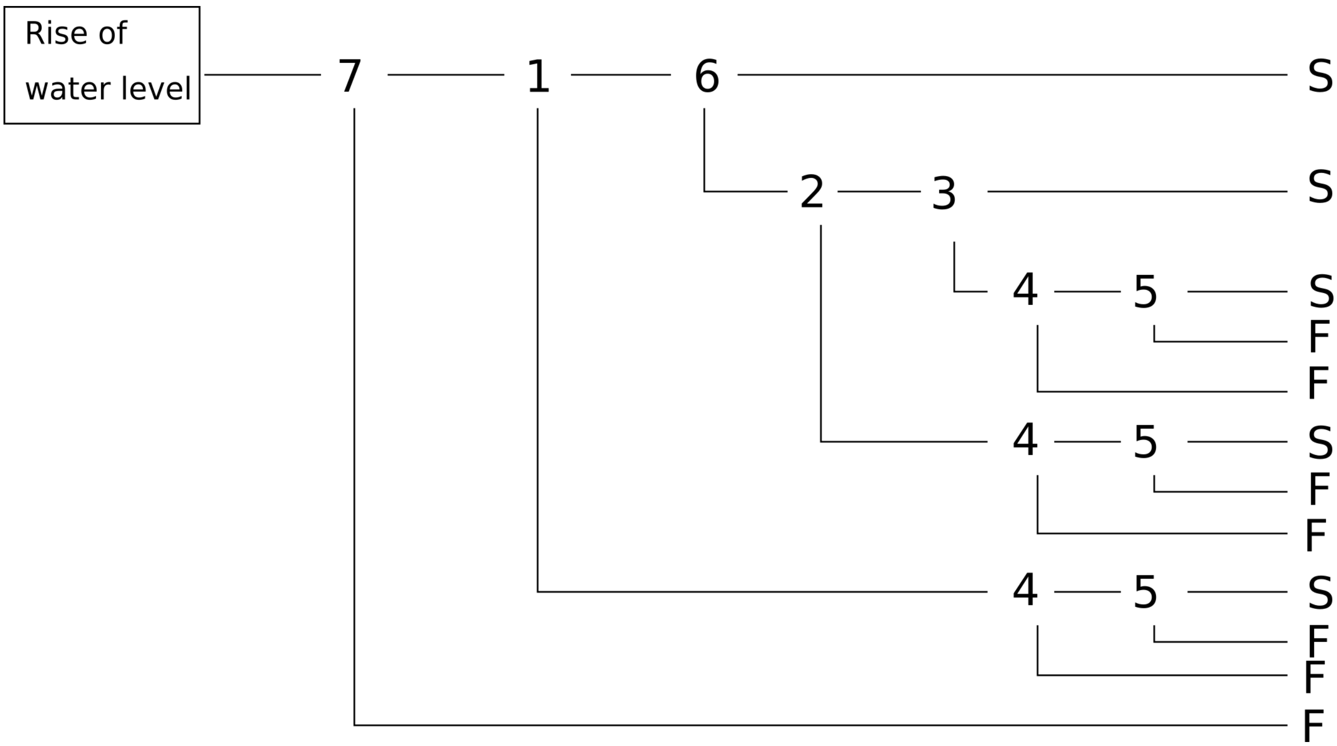

Figure D.5: Representation with an event tree

Failure of the zone is represented by the downward path while not failure of the zone is the straight line. S represent a successful path and F a failure path.