

An Imprecise Bayesian Approach to Thermal Runaway Probability

Marc Fischer, Alexis Vignes

▶ **To cite this version:**

Marc Fischer, Alexis Vignes. An Imprecise Bayesian Approach to Thermal Runaway Probability. 12. International Symposium on Imprecise Probability: Theories and Applications (ISIPTA 2021), Jul 2021, Grenade, Spain. ineris-03268851

HAL Id: ineris-03268851
https://ineris.hal.science/ineris-03268851

Submitted on 23 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Imprecise Bayesian Approach to Thermal Runaway Probability

Marc Fischer MARC.FISCHER@EMSE.FR
Mines de Saint-Étienne, CNRS, UMR 5307 LGF, Centre SPIN, F-42023 Saint-Étienne, Université de Lyon, France

Alexis Vignes ALEXIS.VIGNES@INERIS.FR
INERIS, Parc Technologique ALATA, B.P. 2, F-60550 Verneuil-en-Halatte, France

Abstract

In this pioneering work, an assessment of thermal runaway probability based on simplified *chemical kinetics* has been performed with imprecise Bayesian methods relying on several priors. The physical phenomenon is governed by two chemical kinetic parameters A and Ea. We suppose that their values are considerably uncertain but also that we know the experimental profiles of a chemical species corresponding to their true values, thereby allowing us to compute likelihoods and posteriors corresponding to different levels of information. We are interested in the critical delay time tc beyond which an explosion will certainly occur. The use of several priors allows us to see when the data truly dominate the prior with respect to the probability distribution of tc. It does not appear possible to do so in an orthodox precise Bayesian framework that reduces all forms of uncertainty to a single probability distribution.

Keywords: Robust Bayesianism, explosions, chemical kinetics, principle of indifference

1. Introduction

Explosions remain a very serious threat in the industrial world (Atrkar and Jabbari, 2013; Skob et al., 2020; Ahmed et al., 2012). From a general point of view, an explosion can be defined as a sudden increase in pressure and temperature stemming from an oxidation or other exothermic reactions. They are a complex phenomenon emerging from the interplay of chemistry, heat transfers and fluid dynamics. To make predictions, it is possible to rely either on phenomenological approaches (Proust, 2005) or on CFD (Computational Fluid Dynamics) tools (Ferrara et al., 2006). These CFD tools require a fundamental knowledge of the combustion characteristics of the substances. However, these characteristics are often highly uncertain because of the lack of relevant experimental data and also the absence of standardised experimental approaches to determining fundamental parameters (e.g. laminar flame speeds, oxidation kinetics at high temperatures). Another difficulty which arises when assessing explosion probabilities consists of the factors the thermal runaway process depends upon.

In order to estimate the probability of an explosion in a given situation, we must take into account both the probability distributions of the initial and boundary conditions (aleatory uncertainty) and the uncertainties of the physical and chemical parameters of the model (epistemic uncertainty). For that sake, classical (precise) Bayesian methods are increasingly being employed but they suffer from their inability to properly consider the difference between aleatory and epistemic uncertainty (Schöbi and Sudret, 2019; Ferson and Oberkampf, 2009; Mathon et al., 2010). Routinely, explosion hazards are assessed through the determination of explosion safety parameters according to various standards (e.g. EN 1839:2017 (EN et al., 2017), ISO 10156:207 for gases and vapours (Zakel et al., 2019)).

However, when it comes to determining explosion probabilities in a well-defined hazardous scenario by measuring ignition-sensitive parameters (e.g. minimum ignition energy, auto-ignition temperature), it should be noted that the risk and safety engineer can only access values which correspond to a maximised explosion probability. The explosion probability can thus be potentially considerably overestimated. In order to estimate more realistic probabilities, we would either have to perform costly tests or to use our more fundamental knowledge of *oxidation kinetics* from low to high temperatures. The advantage of such a chemical kinetic approach is that it can be applied to a wide range of industrial scenarios (Warnatz et al., 2017; Peters and Rogg, 1993).

Nevertheless, the chemical kinetic parameters are then the main source of epistemic uncertainty in explosion hazard assessments. They are generally unknown but constrained by physical bounds and a set of experimental data the complete model must be able to reproduce well enough (Shen et al., 2017; Fischer, 2019).

Up until now, most authors who sought to estimate explosion probabilities did not consider the chemical kinetic parameter uncertainties but designed their approach at a macro-level. Ronza et al. (Ronza et al., 2007) and Moosemiller (Moosemiller, 2011) used event trees based on historical data and expert knowledge to predict explosion probabilities. Many researchers apply Bayesian networks or fuzzy Bayesian networks to the assessment of explosion hazard by eliciting expert knowledge or by using the known

frequencies of relevant incidents (Li et al., 2020; Huang et al., 2017; Tong et al., 2018; Yazdi and Kabir, 2017; Lu et al., 2020) Much more works devoted to explosion hazard could be cited but to the best of our knowledge, there does not seem to be any article investigating the *information transfer* of chemical kinetic parameter uncertainties into explosion probabilities.

In this article, we aimed at comparing **precise and imprecise** Bayesian methods for estimating the probability of an explosion while using a highly simplified and idealised model, the thermal theory of explosion of Semenov that involves two chemical kinetic parameters (Semenov, 1942; Shouman, 2006). We created a situation that naturally combines aleatory uncertainty (regarding the initial temperature after some incident) and epistemic uncertainty (concerning the values of the pre-exponential factor A and activation energy Ea of the simplified model that will be presented in the next section).

If the only thing we know about an unknown parameter is that it belongs to a given interval, classical Bayesians usually attribute it a uniform probability density distribution, based on the (in)famous principle of indifference (POI) (Hájek, 2002). The POI stipulates that if we are completely ignorant about which one of n maximally basic possibilities is true, we are rationally compelled to assign each one of them the same probability $1/n$. The problem of the POI is that it can all too easily lead practitioners to mistake ignorance for knowledge. The epistemic state of someone who tossed a coin in a casino 10,000 times and saw its frequency of landing heads closely oscillate around 0.50 should be the same as someone who just discovered the coin in the casino and has no idea whatsoever about whether it is biased or unbiased: both must believe that p(heads) = 0.50 *with the same degree of confidence*. According to the late Bayesian philosopher of science Wesley Salmon, the principle of indifference amounts to magical thinking (Salmon, 1967):

> Knowledge of probabilities is concrete knowledge about occurrences; otherwise it is useless for prediction and action. According to the principle of indifference, this kind of knowledge can result immediately from our ignorance of reasons to regard one occurrence as more probable than another. This is epistemological magic. Of course, there are ways of transforming ignorance into knowledge – by further investigation and the accumulation of more information. It is the same with all "magic": to get the rabbit out of the hat you first have to put him in. The principle of indifference tries to perform "real magic".

Applied to a continuous variable such as a pre-exponential factor A, the POI orders us to use a uniform prior $f_0(A)$. However, since we are equally ignorant about $1/A$, $log10(A)$ and $1/log10(A)$, and many other deterministic functions of A, we should also use priors which are flat with respect to these variables and highly non-uniform with respect to A. In their excellent work devoted to chemical parameter uncertainty propagation, (Frenklach et al., 2007) approvingly stated statisticians Box and Hunter's recommendation for defining priors in chemical kinetics:

> In considering a parameter like the specific rate [constant] φ which is essentially positive, it is probably most realistic to take $\theta = ln\varphi, -\infty \leq \theta \leq \infty$, as locally uniform a priory. This would mean, for example, that having guessed a value of φ, an experimenter would be about equally prepared to accept a value twice as big as he would to accept a value one-half as big.

At first glance, this advice would sound quite reasonable to most chemical kineticists. However, if the authors had also mentioned some logical implications of their approach such as: *"In the absence of any kind of knowledge about φ except its positivity, we should all feel completely confident that $p(1 \leq \varphi \leq 100) \approx 467.51 p(10,001 \leq \varphi \leq 10,100)$, even though both intervals have the same length."*, most practitioners would probably find that rule very strange and arbitrary. If we only know that $\varphi > 0$, how on earth can we deduce such a highly specific result?

The fundamental problem that uniform priors are no longer uniform upon reparametrisation has led Jeffreys to define a standard prior based on the Fisher information that remains the same for any other parameter that is a deterministic function of the first one (Kass and Wasserman, 1996). However, Jeffreys' prior is usually bound to be (highly) non-uniform with respect to many parametrisations of the problem and it thus also illegitimately creates specific knowledge out of ignorance.

To overcome this problem, the field of imprecise Bayesianism (also called robust Bayesianism or Bayesian sensitivity analysis) chooses to describe genuine ignorance through a wide variety of priors that results in a wide variety of posteriors (Walley, 2000; Berger et al., 1994; Insua and Ruggeri, 2012). This approach is being employed in an increasing number of fields such as insurance risks (Boratyńska, 2006), climate science (Tomassini et al., 2007), cybersecurity (Hallgren and Turcotte, 2020), and clinical trials (Greenhouse and Waserman, 1995), to name but a few. However, as Fischer pointed out (Fischer, 2019), chemical kineticists almost always only use one single uniform prior and do not feel concerned about the fact that their posterior might not be data-dominated. As Kass and Wasserman pointed out (Kass and Wasserman, 1996), the reliance on a single so-called uninformative prior is particularly dangerous if the experimental data available are insufficient to constrain the values of the model parameters, and Fischer showed an example where relying on only one flat prior would lead one to reject the model closest to the measurements at hand whereas the use of several priors reveals that this is a spurious and invalid conclusion (Fischer, 2019).

In this paper, we want to compare the use of classical precise Bayesian methods with an imprecise Bayesian approach relying on six different priors for assessing a thermal runaway probability. In Section 2, our methodology is explained. In Section 3, our results are presented and discussed. The article ends with a conclusion and the outlook in Section 4.

2. Methodology

2.1. Computation of the Delay Time Distribution

Our detailed method can be read in Appendix A: Delay time distributions. We first designed a risk scenario. We consider a container filled with gaseous propane C_3H_8 in a closed room with a constant volume initially at atmospheric pressure. We suppose that the room is adiabatic, which means that we neglect the heat losses. Through a leak, a stoichiometric mixture of air-propane is formed in the room. Because of some incident, the ambient temperature jumps from 293.15 K (20 °C) to an initial temperature following a normal distribution: $T_0 \sim N(\mu_{T_0} = 524\ K, \sigma_{T_0} = 20\ K)$. We have $p(T_0 < 440\ K) <$ 1E-04 and $p(T_0 > 600\ K) <$ 1E-04. As a consequence, we can always consider that the initial temperature belongs to the interval $[T_{0,min} = 440\ \text{K}; T_{0,max} = 600\ \text{K}\ ]$ when developing an approximation formula for the delay time. We decided to describe the system through the theory of thermal explosion of (Semenov, 1942). According to it, the reaction progress leads to an increase in temperature which itself spawns an increase in the reaction rate that causes the temperature to rise ever faster (thermal runaway). In our case, the system can be described by the following equations: $C_3H_8 + 5O_2 \rightarrow 3CO_2 + 4H_2O$, $\rho c_v \frac{\partial T}{\partial t} = (-\Delta u_m) r$, $rate = -\frac{d[C_3H_8]}{dt} = -\frac{1}{5}\frac{d[O_2]}{dt} = Ae^{-\frac{Ea}{RT}}[C_3H_8]^a[O_2]^b$.

$[C_3H_8]$ and $[O_2]$ (mol/cm^3) are the concentrations of propane and oxygen, respectively, T(K) is the temperature, $\rho(T)$ (kg/cm^3) is the volumetric mass density of the mixture, c_v (kcal/(kg·K)) is its thermal capacity at constant volume, t(s) is the time, Δu_m (kcal/mol) is the molar reaction energy, $rate$ (mol/cm3/s) is the reaction rate, A(mol,cm,s) is the pre-exponential factor, Ea (kcal/mol) is the activation energy, R = 1.987E-03 kcal/(K mol) is the ideal gas constant, and a and b are reactant coefficients. According to (Westbrook and Dryer, 1981), the following parameter values can be employed: A_0 = 8.6E+11, Ea_0 = 30.0, a = 0.1, and b = 1.65. They used laminar flame speeds to calibrate them.

Further in the study, we shall assume that $A \in$ [6.0E+11 ; 8.0E+13] (mol, cm,s) and $Ea \in [27;46]$ kcal/mol. This corresponds to a situation of strong ignorance in chemical kinetics. We consider an undiluted stoichiometric mixture at atmospheric pressure (p = 1 atm = 101325 Pa), which means we have initially the following mole fractions:
$X_{C_3H_8,0} = \frac{[C_3H_8]_0}{[All\ species]}$ = 0.04057824, $X_{O_2,0} = 0.2028912$, $X_{N_2,0}$ = 0.7565306. The evolution of the variables with time was predicted with the chemical kinetic software Cantera (Goodwin, 2002). We are interested in how quickly the mixture reaches the critical temperature (also called ignition temperature in a technical context) Tc = 766 K beyond which it would be impossible for a technician to intervene to stop the explosion (Reed, 1986). Let tc be the critical delay time defined as $T(tc) = Tc$. If that time is elapsed, an explosion will surely occur. As explained in Appendix A: Delay time distributions, for given values of A and Ea, $log10(tc)$ can be very well approximated by a linear function of $1/T_0$:

$$log10\big(tc_{A,Ea}(T_0)\big) \approx a_{A,Ea}\frac{1}{T_0} + b_{A,Ea} \tag{1}$$

with

$$a_{A,Ea} = \frac{log10\big(tc_{A,Ea}(T_{0,min})\big) - log10\big(tc_{A,Ea}(T_{0,max})\big)}{1/T_{0,min} - 1/T_{0,max}} \tag{2}$$

and

$$b_{A,Ea} = log10\big(tc_{A,Ea}(T_{0,min})\big) - a_{A,Ea}\frac{1}{T_{0,min}}. \tag{3}$$

$log10\big(tc_{A,Ea}(T_{0,min})\big)$ and $log10\big(tc_{A,Ea}(T_{0,max})\big)$ are themselves to a large extent bilinear functions of $log10(A)$ and Ea so that they can be well approximated by a piecewise bilinear interpolation. For that sake, $log10\big(tc_{A,Ea}(T_{0,min})\big)$ and $log10\big(tc_{A,Ea}(T_{0,max})\big)$ were computed for 30*30 values of (A, Ea) numerically with Cantera in order to obtain a training set. For given values of A, Ea and T_0, we can avoid calling Cantera by estimating $log10\big(tc_{A,Ea}(T_{0,min})\big)$ and $log10\big(tc_{A,Ea}(T_{0,max})\big)$ through the piecewise bilinear interpolation and then approximate $log10\big(tc_{A,Ea}(T_0)\big) \approx a_{A,Ea}\frac{1}{T_0} + b_{A,Ea}$ through the linear interpolation formula. The quality of the model has been tested by generating 3000 random values of (A, Ea) and T_0 with $A \sim U(A_{min}, A_{max})$, $Ea \sim U(Ea_{min}, Ea_{max})$ and $T_0 \sim N(\mu_{T_0}, \sigma_{T_0})$. We systematically computed $tc_{A,Ea}(T_0)$ with Cantera and $tc_{pred,A,Ea}(T_0)$ by interpolation and then the relative difference r

$$r = \frac{tc_{A,Ea}(T_0) - tc_{pred,A,Ea}(T_0)}{min\big(tc_{A,Ea}(T_0), tc_{pred,A,Ea}(T_0)\big)}. \tag{4}$$

For more than 99.3% of the points, we have $r \leq 5\%$. We have $max(r) = 12.32\%$ which is reached for a very high and utterly unproblematic delay time superior to 2E+08 s. Given the fact that we are not interested in numerical accuracy but in understanding the behaviour of probabilistic approaches to thermal runaway risks, we deemed that level of error to be acceptable. As demonstrated in Appendix A,

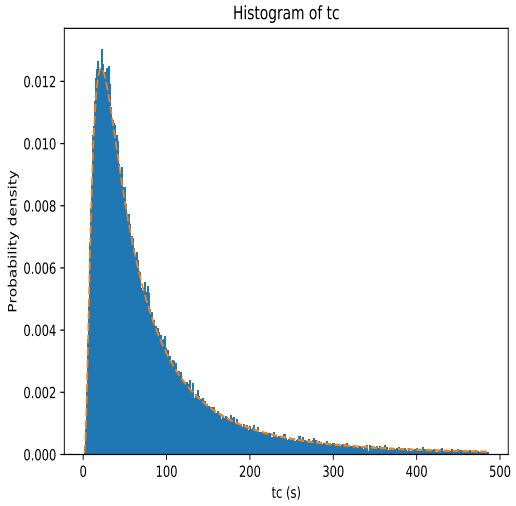


Figure 1: $f(tc)\,for\,A = A_0 = 8.6E+11$ and $Ea = Ea_0 = 30kCal/mol$

for given A and Ea, the probability density of tc is given by

$$f(tc|A,Ea) = \frac{a_{A,Ea} ln(10)}{tc\left(b_{A,Ea} ln(10) - ln(tc)\right)^2}$$
$$\phi\left(\frac{a_{A,Ea}}{log10(tc) - b_{A,Ea}}, \mu_{T_0}, \sigma_{T_0}\right).$$

This analytical formula has been compared with an empirical distribution obtained by generating 10,000 values of $T_0 \sim N(\mu_{T_0} = 524K, \sigma_{T_0} = 20K)$ and computing $tc_{A,Ea}(T_0)$ by using the linear approximation. As can be seen in Figure 1, the analytical distribution of tc corresponds very well to the empirical one.

2.2. Priors

Chemical kineticists often define the prior in such a way that A and Ea are stochastically independent (which might be a problematic assumption, see (Held et al., 2008)) and that it is uniform with respect to $log10(A)$ and Ea (Plessis, 2013; Hsu et al., 2009; Huan and Marzouk, 2013). Consequently, we defined our first prior $f_{0,1}\left(log10(A), Ea\right)$ as uniform with respect to $log10(A)$ and Ea. We then defined our second prior $f_{0,2}$ in such a way that it is uniform with respect to A and Ea. Our third prior $f_{0,3}$ is uniform with respect to $1/log10(A)$ and Ea. The fourth prior $f_{0,4}$ is uniform with respect to $1/log10(A)$ and $1/Ea$. The fifth prior $f_{0,5}$ is uniform with respect to $1/A$ and Ea. The sixth prior $f_{0,6}$ is uniform with respect to A and $1/Ea$. For the sake of the present study, we ignored the problem of the stochastic independence of $log10(A)$ and Ea but intend to explore this issue in future works. The prior distributions along with their analytical expressions can be seen in Appendix B. It is a well-known problem of robust Bayesian analysis that the set of priors must be large enough to avoid the introduction of spurious information that could dominate the experimental data (Walley, 2000) but not so wide as to lead to nearly vacuous posteriors that would make it very hard to draw any practical conclusion (Held et al., 2008).

From a pragmatic and pedagogical point of view, these six priors are a good way to present an imprecise framework to chemical kineticists as all priors are uniform with respect to some reformulation of the kinetic parameters so that it would be arbitrary to only rely on the results derived from one such prior.

The elicitation of priors through subjective means Daneshkhah et al. (2017) is outside the scope of the present study.

2.3. Measurements and Epistemic Situations

We created "experimental" data allowing us to determine the posterior probability distributions of $log10(A)$ and Ea. We considered a constant-volume adiabatic reactor at atmospheric pressure with a very diluted mixture of propane and oxygen: X_{C3H8} = 1E-05, X_{O2} = 5E-05 and X_{N2} = 0.99994 so that the temperature $T(t)$ remains nearly constant. Using the "true" values A_0 = 8.60E+11 and Ea_0 = 30.00 kcal/mol, we generated mole fraction profiles of propane at different temperatures. We then randomly chose several time points and generated normally distributed noise in such a way that $X_{C_3H_8,exp}(t_j) = X_{C_3H_8}(A_0, Ea_0, t_j) + \varepsilon_{j,A_0,Ea_0}$ and $\varepsilon_{j,A_0,Ea_0} \sim N\left(0, \sigma_{j,A_0,Ea_0}\right)$ with the standard deviation $\sigma_{j,A_0,Ea_0} = \sigma_r X_{C3H8}(A_0, Ea_0, t_j)$ where σ_r is the relative standard deviation that always remains constant during an experiment. In practice, the true parameters A_0 and Ea_0 are of course unknown. For each measurement j of a given experiment, we use the expression of the standard deviation σ_{j,A_0,Ea_0} defined just above.

As explained in Appendix C, we distinguished four epistemic situations:

- A: we only know that $A \in$ [6.0E+11 ; 8.0E+13] (mol, cm,s) and $Ea \in [27;46]$ kcal/mol.
- B: we have one profile of X_{C3H8} with 6 time points measured at 1845 K with $\sigma_r = 25\%$.
- C: We have two profiles of X_{C3H8} with 6 time points measured at 1135 K and 2249 K with $\sigma_r = 25\%$.
- D: We have four profiles of X_{C3H8} with 10 time points measured at 1135 K, 1478 K, 1845 K, and 2249 K with $\sigma_r = 6\%$.

As an example, X_{C3H8} at 1845 K with $\sigma_r = 25\%$ can be seen in Figure 2. The other "measurements" are shown in Appendix C. The log-likelihood can be expressed as

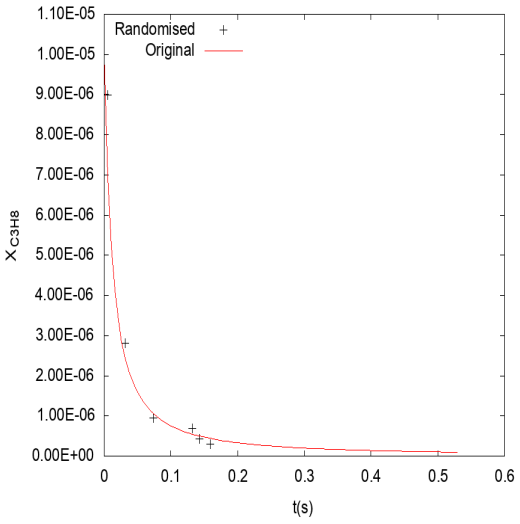


Figure 2: X_{C3H8} at 1845 K with $\sigma_r = 25\%$

follows:

$$l(data|log10(A),Ea) = \sum_{i=1}^{m}\sum_{j=1}^{n_{t,i}}\left(-\frac{1}{2}ln\left(2\pi\sigma_{i,j}^2\right) - \frac{1}{2\sigma_{i,j}^2}\left(X_{C3H8,i}(t_j,A,Ea) - X_{C3H8,exp,i}(t_j)\right)^2\right) \quad (5)$$

whereby m is the number of experiments and $n_{t,i}$ is the number of time points for experiment i and $\sigma_{i,j} = \sigma_r X_{C_3H_8,i,A_0,Ea_0}(t_j)$ as explained above. For a given prior $f_0\left(log10(A),Ea\right)$ and a set of experimental data, the **joint posterior probability density** of $log10(A)$ and Ea can be expressed by Eq. 6

$$f\left(log10(A),Ea|Data\right) = \frac{L(Data|log10(A),Ea)f_0\left(log10(A),Ea\right)}{\iint_{log10(A),Ea} L(Data|log10(A),Ea)f_0\left(log10(A),Ea\right)dlog10(A)dEa} \quad (6)$$

The integration was carried out with a first-order Euler explicit method (Hoffman and Frankel, 2018). The posteriors in Situation B, C and D (see 2.3) obtained with the first prior and 300*300 values of (A, Ea) are shown in Figure 3, 4 and 5, respectively. It is very obvious that there was a strong reduction in uncertainty between Situation B (involving one imprecise experiment at one temperature) and Situation D (involving four precise experiments at four temperatures). The other posterior probability densities of $log10(A)$ and Ea are displayed as contour-plots in fischer21c-supp.

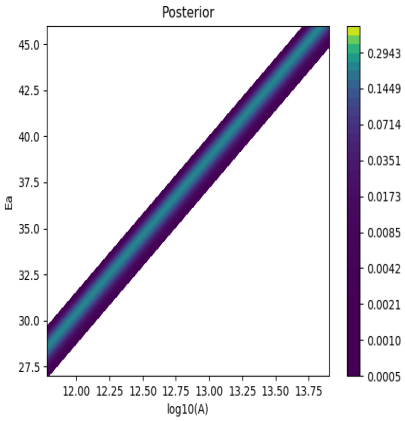


Figure 3: $f_{1,B}\left(log10(A),Ea\right)$

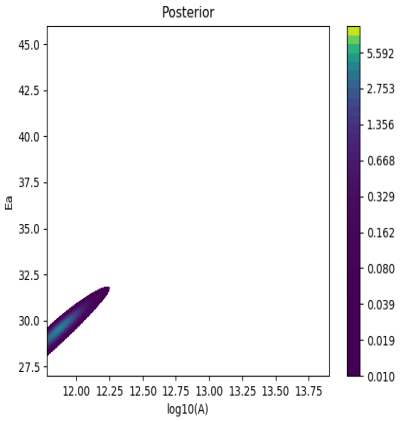


Figure 4: $f_{1,C}\left(log10(A),Ea\right)$

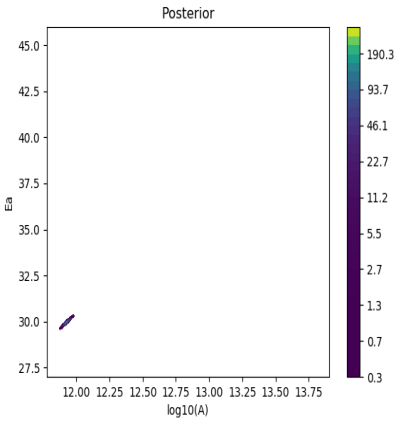


Figure 5: $f_{1,D}\left(log10(A),Ea\right)$

Prior	tc_{25}	tc_{50}	tc_{75}	$p_{critical}$
$f_{0,1}$	34.29	2801.66	2.34E+05	0.2424
$f_{0,2}$	7.85	649.06	5.45E+04	0.3260
$f_{0,3}$	39.22	3201.74	2.68E+05	0.2348
$f_{0,4}$	9.39	356.02	4.00E+04	0.3360
$f_{0,5}$	149.45	1.21E+04	9.99E+05	0.1585
$f_{0,6}$	1.80	69.51	8019.06	0.4476
$\mathbf{A_0, Ea_0}$	**28.84**	**55.04**	**108.68**	**0.2636**

Table 1: Features of $f(tc)$ in situation A (times in s).

Prior	tc_{25}	tc_{50}	tc_{75}	$p_{critical}$
$f_{0,1}$	322.64	5667.50	1.02E+05	0.0600
$f_{0,2}$	4.61E+04	2.67E+05	1.19E+06	3.4E-03
$f_{0,3}$	234.54	3630.43	7.18E+04	0.0690
$f_{0,4}$	108.43	1035.65	2.15E+04	0.1018
$f_{0,5}$	35.05	114.69	604.56	0.2183
$f_{0,6}$	2.60E+04	1.93E+05	9.59E+05	6.9E-03
$\mathbf{A_0, Ea_0}$	**28.84**	**55.04**	**108.68**	**0.2636**

Table 2: Features of $f(tc)$ in situation B (times in s).

2.4. Computation of $f(tc)$ as a Function of $f\big(log10(A), Ea\big)$

For a joint probability density of $log10(A)$ and Ea $f\big(log10(A), Ea\big)$ (which could be either a prior or a posterior), the probability density of tc is given by Eq 7.

$$f(tc) = \int_{log10(A)\in[log10(A_{min}),log10(A_{max})],Ea\in[Ea_{min},Ea_{max}]} f\big(tc|log10(A), Ea\big) f_0\big(log10(A), Ea\big) dlog10(A) dEa \tag{7}$$

We then numerically computed the three quartiles of tc (tc_{25}, tc_{50}, tc_{75}) and $p_{critical} = p(tc \leq 30s)$ which is the probability that the temperature rises so fast that a technician would not be able to step in. For A_0 and Ea_0, we have tc_{25} = 28.86 s, tc_{50} = 55.08 s, tc_{75} = 108.79 s, and $p_{critical}$ = 0.2634.

3. Results and Discussion

3.1. Estimation of $f(tc)$

For the four situations A, B, C, and D, we computed $f(tc)$ and the four features mentioned in 2.4. All the results can be seen in Appendix D.

In Table 1, the features of $f(tc)$ are shown in situation A where we have no measurements and only know that $A \in$ [6.0E+11 ; 8.0E+13] (mol, cm,s) and $Ea \in [27;46]$ kcal/mol. Our representation of our ignorance through the six priors results in widely different distributions $f(tc)$.

Table 2 shows the features of the six $f(tc)$ in situation B where the six priors have been updated by the information contained in one experiment where $X_{C3H8}(t)$ has been measured at constant temperature (see 2.3). These stark differences show that our knowledge of A and Ea is much too imprecise to draw any practical conclusions and that we need new and better experimental data to sharpen their joint posterior probability distribution and decorrelate the two parameters.

The results for situation C can be seen in Table 3. The posteriors are considerably more similar and closer to the true values obtained with A_0 and Ea_0. However, the delay

Prior	tc_{25}	tc_{50}	tc_{75}	$p_{critical}$
$f_{0,1}$	22.56	44.03	88.39	0.3508
$f_{0,2}$	23.74	46.62	94.41	0.3311
$f_{0,3}$	22.49	43.86	87.99	0.3521
$f_{0,4}$	22.24	43.32	86.77	0.3566
$f_{0,5}$	21.57	41.83	83.22	0.3689
$f_{0,6}$	23.45	46.00	93.02	0.3358
$\mathbf{A_0, Ea_0}$	**28.84**	**55.04**	**108.68**	**0.2636**

Table 3: Features of $f(tc)$ in situation C (times in s).

times are systematically under-predicted whereas $p_{critical}$ is systematically over-predicted. As can be seen in Figure 4, while situation C is a clear improvement over situation B, the parameter values are still considerably uncertain. This is plausibly due to the fact that in Situation C, the highest value of the likelihood function is obtained for $A_{max,C}$ = 7.42E+11 (mol, cm, s) and $Ea_{max,C}$ = 29.54 kcal/mol instead of A_0 = 8.60E+11 (mol, cm, s) and Ea_0 = 30.00 kcal/mol. This leads to a higher reaction rate and thus to shorter ignition delay times and a larger explosion probability which result in tc_{25} = 22.14 s, tc_{50} = 41.48 s , tc_{75} = 81.66 s , $p_{critical}$ = 0.3639 for $A_{max,C}$ and $Ea_{max,C}$. The *systematic* under-prediction of the delay times could possibly be avoided by using a parametric family of prior probability distributions resulting in larger posterior probability intervals.

Finally, the results obtained by updating the priors with the data of four experiments with a higher accuracy are summarised in Table 4. One can see that the discrepancies have become much narrower and that they are very close to the true values of the variables. The small system-

Prior	tc_{25}	tc_{50}	tc_{75}	$p_{critical}$
$f_{0,1}$	28.55	54.53	107.77	0.2672
$f_{0,2}$	28.61	54.65	108.03	0.2664
$f_{0,3}$	28.54	54.52	107.75	0.2672
$f_{0,4}$	28.53	54.49	107.70	0.2674
$f_{0,5}$	28.48	54.40	107.51	0.2680
$f_{0,6}$	28.60	54.62	107.98	0.2666
$\mathbf{A_0, Ea_0}$	**28.84**	**55.04**	**108.68**	**0.2636**

Table 4: Features of $f(tc)$ in situation D (times in s).

atic underprediction of the delay times and overprediction of $p_{critical}$ probably stem from the fact that in Situation D, the highest value of the likelihood function is reached for $A_{max,D} = 8.32\text{E+}11$ (mol, cm, s) and $Ea_{max,D} = 29.92$ kcal/mol instead of $A_0 = 8.60\text{E+}11$ (mol, cm, s) and $Ea_0 = 30.00$ kcal/mol. This results in a higher reaction rate and hence in shorter ignition delay times and a greater value of the explosion probability. For $A_{max,D}$ and $Ea_{max,D}$, we have $tc_{25} = 27.74$ s, $tc_{50} = 52.85$ s , $tc_{75} = 104.15$ s , $p_{critical} = 0.2773$.

3.2. Epistemic Consequences

Precise Bayesianism stipulates that the epistemic state of a rational agent can be perfectly well described through a *single* probability distribution. This is nicely captured by this statement of de Finetti (De Finetti, 1931; Vicig and Seidenfeld, 2012) who is widely regarded as one of the most important Bayesian mathematicians of the 20th century:

> In experimental sciences, the world of feelings is replaced by a fictitious world where quantities have an exactly measurable value; in probability theory, I replace my vague, elusive mood with that of a fictitious agent with no uncertainty in grading the degrees of his beliefs.

According to most precise Bayesian theorists, uncertainty is nothing beyond and above a probability distribution over all possible states of an event or process.

The results obtained during this work challenge this view but also help us see which properties a good imprecise Bayesian analysis ought to have.

The first problem concerns the arbitrariness of the choice of the probability distribution. As can be seen through the stark differences between the distributions of tc, in situation B the prior strongly dominates the likelihood based on the experimental data. It seems completely arbitrary to say that we ought to only consider the results stemming from one prior (such as the one uniform with respect to $log10(A)$ and Ea) and disregard all other ones (such as those uniform with respect to A and Ea or $1/A$ and $1/Ea$). For if we are at the beginning completely ignorant about the values of $log10(A)$ and Ea, we should logically also be completely ignorant about, say, $1/A$ and $1/Ea$. Using Jeffreys' prior which is invariant with respect to reparametrisation would not help, as it would be non-uniform (probably strongly non-uniform) for at least some parametrisations. An easy way to avoid these problems is to represent the epistemic state of an engineer in situation B through *all distributions* at the same time (and possibly also distributions stemming from other priors).

The second problem is the inability of the orthodox precise Bayesian position to distinguish knowledge and ignorance. Let us consider the probability $p_{critical}$ that the thermal runaway gets out of hand. For an imprecise Bayesian, the distinction between knowledge and ignorance can be captured by the differences between the posteriors and the values derived out of them. In situation A where we only know the parameter bounds, $p_{critical,A}$ takes on values between 0.1585 and 0.4476 and we can see that the average delay times t_{25}, t_{50} and t_{75} differ by orders of magnitude. In situation B where we only have one experiment which does not allow us to separate $log10(A)$ and Ea, $p_{critical,B}$ takes on values between 3.4E-03 and 0.2183 and the average delay times again differ by orders of magnitude. In situation D where we have more accurate measurements from four experiments at different temperatures, $p_{critical,D}$ takes on values between 0.2664 and 0.2680, which is close to 0.2634 and the differences between the delay times is always smaller than 0.5 %. This corresponds to a situation of warranted knowledge. The main limitation of the very simple priors we have chosen can be seen in situation C: the relative difference between the delay times is smaller than 12 % but the lowest and highest values of $p_{critical,D}$ (0.3311 and 0.3689) are far from the true value (0.2634). This outcome could plausibly be avoided by using a family of near-ignorance priors in the exponential family (Benavoli and Zaffalon, 2015; Quaeghebeur and De Cooman, 2005).

While the crude priors we considered here do not allow us to discriminate between cases where the distribution of tc is inaccurate, they do permit us to recognise situations where the data truly dominate the priors. Indeed, Bayesian convergence theorems (commonly known as "The priors wash out!") (Hawthorne, 1994) show that the different posteriors are bound to converge towards a singular probability density distribution equal to 1 for $A = A_0$ and $Ea = Ea_0$ with a favourable rate if the chosen priors are not too extreme.

Now, how could a Bayesian who is, for some reason, allergic to the very idea of interval probability manage to capture the distinction between knowledge and ignorance whilst using only one prior distribution? One possible way to do this would be to rely on the difference between "the weight of the argument" and the balance of evidence that was first introduced by British economist John Maynard Keynes (Keynes, 1921) and was recently explored by (Hill, 2019). Let us suppose that this Bayesian chooses to use the prior which is uniform with respect to $log10(A)$ and Ea. In situation A, all pairs of $(log10(A), Ea)$ have the same probability density. This corresponds to a situation of *extreme ignorance* (or maximum ignorance given the parameter bounds). Thus in situation A, we have $p_{critical,A} = 0.2424$ but this value is **extremely unreliable** as the weight of the argument is equal to zero. Likewise, in situation B $p_{critical,B} = 0.06$ but this value is **strongly unreliable** because $log10(A)$ and Ea are strongly correlated as can be visualised in Figure 3. $p_{critical,D} = 0.2672$ and this value is **strongly reliable** as shown by Figure 5 where the likely

values of $(log10(A), Ea)$ occupy a very small region. In situation C, $p_{critical,C} = 0.3508$ but the degree of reliability (or weight of the argument) is weaker as a much larger range of values are equally probable.

The reliability of $p_{critical}$, (Keynes' weight of the argument) would also have obvious consequences for decision-making and actions. In situations D, we reliably know that the probability of an uncontrollable thermal runaway is way too high, which means we must absolutely store the propane in a room where this type of incident raising the initial temperature T_0 is impossible. In situations A, B and C, we do not reliably know $p_{critical}$ so that we need to collect more relevant experimental data in order to strongly narrow down the range of possible values of (A, Ea).

Such a Bayesian approach could also be applied to the example of the coin mentioned in the introduction 1. If we know absolutely nothing about the coin (except that it cannot land on edges), we would believe that $p(heads) = 0.5$ while also knowing that this value is extremely unreliable so that ambiguity-averse individuals would be unwilling to bet any amount of money on either outcome. If we saw the relative frequency of heads during 10,000 tosses oscillate in a very narrow interval around 0.5, we would believe that $p(heads) = 0.5$ and that this value is strongly reliable so that ambiguity-averse individuals would have no problem participating in bets. As shown by (Feduzi, 2010), Keynes wanted both the weight of the argument and the balance of evidence (the precise probability value in his framework) to play a role in decision-making but was unsure about how to achieve this because of the stopping-rule problem.

If applied consistently, this variety of Bayesianism could *in this specific situation* account for the difference between ignorance and knowledge. It would also provide us with an explanation of Ellberg's paradox (Ellsberg, 1961) that completely respects the intuitions of the betting agents. However, it deviates so strongly from the precise Bayesian orthodoxy that it deserves to be considered a form of imprecise probability which fully recognises that probability can have different degrees of reliability.

That being said, this solution (which basically relies on some sorts of second-order probabilities) would not work in other situations. As Walley (1996) showed, in problems involving a multinomial distribution (such as guessing the colour of the next marble drawn from a urn), inferences based on a uniform prior depends on how the possibility space is defined and partioned. This is very problematic, as we do not have *a priory* any more reason to suppose that $p(red) = p(yellow) = p(green) = p(black)$ than to assume that $p(red) = p(non-red)$ or $p(yellow) = p(non-yellow)$ and that we cannot automatically assume that the probability is uniform with respect to the simplest partition Kelly (2011).

4. Conclusion and Outlook

To the best of our knowledge, the evaluation of explosion and more specifically thermal runaway probabilities mostly occurs through the use of Bayesian networks based on the opinions of experts and on frequency data while representing uncertainty through a single probability distribution. What is more, while chemical kinetic parameter values are one main source of epistemic uncertainty, no effort has been made to transfer their uncertainties into predictions of explosion, so far as we know. In this pioneering work, we aimed at filling this gap by conducting a study combining aleatory uncertainty (the values of the initial temperature T_0) and epistemic uncertainty (the values of the parameters A and Ea) to predict the distribution of the critical time tc before the thermal runaway gets completely out of hand.

After an introduction in Section 1, we presented our methodology in Section 2 by detailing the physical system and the underlying differential equations, the computation of the PDF of tc for given values of A and Ea, the choice of the priors, the experiments used to update the priors, and how to compute $f(tc)$ given a joint PDF of (A,Ea). In Section 3, we presented our results along with their interpretation. Situation A corresponds to complete ignorance, Situation B to strong ignorance, situation D to very good knowledge, and situation C to insufficient knowledge. Our imprecise Bayesian method relying on 6 *crude* priors allows us to recognise that we know $p_{critical}$ in situation D but are very ignorant in situation A and B. It is much harder to decide how accurate $p_{critical}$ is in situation C based solely on the differences between the posteriors.

A Bayesian method considering only one prior but also a degree of accuracy for the various probabilities could potentially also capture the crucial distinction between knowledge and ignorance. However, it would strongly deviate from the precise Bayesian orthodoxy and would be a theory of imprecise probability in its own right. It would also fail in more complex situations such as those involving multinomial distributions.

There are several aspects we intend to explore in future works:

- The thermal runaway model we used is a huge simplification of reality. In addition to considering the thermal self-reinforcement of the reaction, the model should also include the chain reactions involving free radicals and the competition between ramification and chain termination reactions that can lead either to an explosion or to the end of the overall reaction (Warnatz et al., 2017). Ultimately, we shall also consider the complex interactions between chemistry, mass transfer, velocity fields, and heat transfer through complex CFD (Computational Fluid Dynamics) simulations (Seok et al., 2013).

- We considered only an homogeneous gas-phase reaction. Considering liquid-gas or solid-gas explosions would be very relevant for safety engineers as this type of scenario is more likely to be encountered in the industrial world. One example is dust explosion which can have disastrous consequences (Eckhoff, 2005). The simulation of such heterogeneous reactions would be computationally much more demanding (Murillo et al., 2013; Ermoline et al., 2013; Williams, 1979).
- The uniform priors we considered are very simple and as we saw, they are not good at identifying the level of inaccuracy in cases of partial ignorance (situation C). It would be interesting to find out whether our approach would be better at distinguishing different degrees of ignorance by using a class of Gaussian priors also characterised by different coefficients of correlation. However, it does not appear possible to obtain analytical versions of the posteriors, especially not if we use more realistic and complex models. In complex situations involving many parameters, we would then be left with no other choice than to rely on the MCMC (Markov-chain Monte-Carlo) algorithm to approximate the posteriors. Given the very long duration of each CFD simulation, we would also need to develop new surrogate models that are sufficiently trustworthy over a wide range of conditions.

References

Mebarki Ahmed, Sandra Jerez, Igor Matasic, Gaëtan Prodhomme, and Mathieu Reimeringer. Explosions and structural fragments as industrial hazard: domino effect and risks. *Procedia Engineering*, 45:159–166, 2012.

Roshan Sedigheh Atrkar and Gharedagh Moosa Jabbari. Economic consequence analysis of fire and explosion in petrochemical feed and product pipelines network. 2013.

Alessio Benavoli and Marco Zaffalon. Prior near ignorance for inferences in the k-parameter exponential family. *Statistics*, 49(5):1104–1140, 2015.

James O Berger, Elías Moreno, Luis Raul Pericchi, M Jesús Bayarri, José M Bernardo, Juan A Cano, Julián De la Horra, Jacinto Martín, David Ríos-Insúa, Bruno Betrò, et al. An overview of robust bayesian analysis. *Test*, 3 (1):5–124, 1994.

Agata Boratyńska. Robust bayesian prediction with asymmetric loss function in poisson model of insurance risk. 2006.

Alireza Daneshkhah, Amin Hosseinian-Far, Tabassom Sedighi, and Maryam Farsi. Prior elicitation and evaluation of imprecise judgements for bayesian analysis of system reliability. In *Strategic Engineering for Cloud Computing and Big Data Analytics*, pages 63–79. Springer, 2017.

Bruno De Finetti. Sul significato soggettivo della probabilita. *Fundamenta mathematicae*, 17(1):298–329, 1931.

Rolf Eckhoff. Current status and expected future trends in dust explosion research. *Journal of loss prevention in the process industries*, 18(4-6):225–237, 2005.

Daniel Ellsberg. Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, pages 643–669, 1961.

BS EN et al. Determination of the explosion limits and the limiting oxygen concentration (loc) for flammable gases and vapours. *British Standards Institution*, 2017.

Alexandre Ermoline, Deniz Yildiz, and Edward L Dreizin. Model of heterogeneous combustion of small particles. *Combustion and flame*, 160(12):2982–2989, 2013.

Alberto Feduzi. On keynes's conception of the weight of evidence. *Journal of Economic Behavior & Organization*, 76(2):338–351, 2010.

Gabriele Ferrara, Almerinda Di Benedetto, Ernesto Salzano, and Gennaro Russo. Cfd analysis of gas explosions vented through relief pipes. *Journal of Hazardous Materials*, 137(2):654–665, 2006.

Scott Ferson and William L Oberkampf. Validation of imprecise probability models. *International Journal of Reliability and Safety*, 3(1-3):3–22, 2009.

Marc Fischer. On the usefulness of imprecise bayesianism in chemical kinetics. In *International Symposium on Imprecise Probabilities: Theories and Applications*, pages 203–215, 2019.

Michael Frenklach, Andrew Packard, and Ryan Feeley. Optimization of reaction models with solution mapping. *Comprehensive Chemical Kinetics*, 42:243–291, 2007.

David G Goodwin. Cantera c++ user's guide. *California Institute of Technology*, 2002.

Joel B Greenhouse and Larry Waserman. Robust bayesian methods for monitoring clinical trials. *Statistics in medicine*, 14(12):1379–1391, 1995.

Alan Hájek. Interpretations of probability. 2002.

Karl Lars Yvon Hallgren and Melissa Turcotte. Robust bayesian change detection for cyber-security applications. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2020.

James Hawthorne. On the nature of bayesian convergence. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, volume 1994, pages 241–249. Philosophy of Science Association, 1994.

Hermann Held, Thomas Augustin, and Elmar Kriegler. Bayesian learning for a class of priors with prescribed marginals. *International journal of approximate reasoning*, 49(1):212–233, 2008.

Brian Hill. Confidence in belief, weight of evidence and uncertainty reporting. In *International Symposium on Imprecise Probabilities: Theories and Applications*, pages 235–245, 2019.

Joe D Hoffman and Steven Frankel. *Numerical methods for engineers and scientists*. CRC press, 2018.

Shuo-Huan Hsu, Stephen D Stamatis, James M Caruthers, W Nicholas Delgass, Venkat Venkatasubramanian, Gary E Blau, Mike Lasinski, and Seza Orcun. Bayesian framework for building kinetic models of catalytic systems. *Industrial & Engineering Chemistry Research*, 48 (10):4768–4790, 2009.

Xun Huan and Youssef M Marzouk. Simulation-based optimal bayesian experimental design for nonlinear systems. *Journal of Computational Physics*, 232(1):288–317, 2013.

Yimiao Huang, Guowei Ma, and Jingde Li. Grid-based risk mapping for gas explosion accidents by using bayesian network method. *Journal of Loss Prevention in the Process Industries*, 48:223–232, 2017.

David Ríos Insua and Fabrizio Ruggeri. *Robust Bayesian Analysis*, volume 152. Springer Science & Business Media, 2012.

Robert E Kass and Larry Wasserman. The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91(435):1343–1370, 1996.

Kevin T Kelly. Simplicity, truth, and probability. In *Philosophy of statistics*, pages 983–1024. Elsevier, 2011.

John Maynard Keynes. *A treatise on probability*. Macmillan and Company, limited, 1921.

Min Li, Hetang Wang, Deming Wang, Zhenlu Shao, and Shan He. Risk assessment of gas explosion in coal mines based on fuzzy ahp and bayesian network. *Process Safety and Environmental Protection*, 135:207–218, 2020.

Yunmeng Lu, Tiantian Wang, and Tiezhong Liu. Bayesian network-based risk analysis of chemical plant explosion accidents. *International journal of environmental research and public health*, 17(15):5364, 2020.

Bree R Mathon, Metin M Ozbek, and George F Pinder. Dempster–shafer theory applied to uncertainty surrounding permeability. *Mathematical Geosciences*, 42(3):293–307, 2010.

Mike Moosemiller. Development of algorithms for predicting ignition probabilities and explosion frequencies. *Journal of Loss Prevention in the Process Industries*, 24 (3):259–265, 2011.

Carlos Murillo, Olivier Dufaud, Omar Lopez, Laurent Perrin, Alexis Vignes, and Felipe Munoz. Cfd modelling of nanoparticles dispersion in a dust explosion apparatus. In *14. International Symposium on Loss Prevention and Safety Promotion in the Process Industry*, volume 31, pages 889–894. AIDIC. Milano, 2013.

Norbert Peters and Bernd Rogg. *Reduced Kinetic Mechanisms for Applications in Combustion Systems*. Springer edition, 1993.

Sylvain Plessis. Calibration of the CO + OH -> CO2 + H rate constant with direct and indirect information. 8th US National Combustion Meeting, 2013.

Christophe Proust. The usefulness of phenomenological tools to simulate the consequences of dust explosions. the experience of effex. In *International ESMG Symposium" Process safety and industrial explosion protection"*, page NC. ESMG. Hamm, 2005.

Erik Quaeghebeur and Gert De Cooman. Imprecise probability models for inference in exponential families. In *4th International Symposium on Imprecise Probabilities and Their Applications*, pages 287–296. International Society for Imprecise Probability: Theories and Applications . . . , 2005.

Richard J Reed. North american combustion handbook, vol. 1: Combustion, fuels, stoichiometry. *Heat transfer, Fluid flow*, 1986.

A Ronza, JA Vilchez, and Joaquim Casal. Using transportation accident databases to investigate ignition and explosion probabilities of flammable spills. *Journal of hazardous materials*, 146(1-2):106–123, 2007.

Wesley Salmon. *The Foundations of Scientific Inference*. University of Pittsburgh, 1967.

Roland Schöbi and Bruno Sudret. Global sensitivity analysis in the context of imprecise probabilities (p-boxes) using sparse polynomial chaos expansions. *Reliability Engineering & System Safety*, 187:129–141, 2019.

Nikolai Nikolaevich Semenov. Thermal theory of combustion and explosion. 3; theory of normal flame propagation. 1942.

Jun Seok, Se-Min Jeong, Jong-Chun Park, and Jeom-Kee Paik. Cfd simulation of methane combustion for estimation of fire and explosion in offshore plant. *Journal of Ocean Engineering and Technology*, 27(2):59–68, 2013.

Xiaobo Shen, Bo Zhang, Xiaoliang Zhang, and Guangli Xiu. Explosion characteristics of methane-ethane mixtures in air. *Journal of Loss Prevention in the Process Industries*, 45:102–107, 2017.

Ahmad R Shouman. A review of one aspect of the thermal-explosion theory. *Journal of engineering mathematics*, 56(2):179–184, 2006.

YA Skob, ML Ugryumov, and EA Granovskiy. Numerical assessment of hydrogen explosion consequences in a mine tunnel. *International Journal of Hydrogen Energy*, 2020.

Lorenzo Tomassini, Peter Reichert, Reto Knutti, Thomas F Stocker, and Mark E Borsuk. Robust bayesian uncertainty analysis of climate system properties using markov chain monte carlo methods. *Journal of Climate*, 20(7): 1239–1254, 2007.

Xing Tong, Weipeng Fang, Shuaiqi Yuan, Jinyu Ma, and Yiping Bai. Application of bayesian approach to the assessment of mine gas explosion. *Journal of Loss Prevention in the Process Industries*, 54:238–245, 2018.

Paolo Vicig and Teddy Seidenfeld. Bruno de finetti and imprecision: Imprecise probability does not exist! *International Journal of Approximate Reasoning*, 53(8): 1115–1123, 2012.

Peter Walley. Inferences from multinomial data: learning about a bag of marbles. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):3–34, 1996.

Peter Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24:125–148, 2000.

Juergen Warnatz, Ulrich Maas, and Robert W Dibble. Combustion: physical and chemical fundamentals, modeling and simulation, experiments, pollutant formation. 2006. *Int. J. Comp. Meth. and Exp. Meas*, 5(1), 2017.

Charles K. Westbrook and Frederick Dryer. Simplified reaction mechanisms for the oxidation of hydrocarbon fuels in flames. *Combustion science and technology*, 27 (1–2):31–43, 1981.

Alan Williams. Fundamentals of oil combustion. In *Energy and Combustion Science*, pages 135–147. Elsevier, 1979.

Mohammad Yazdi and Sohag Kabir. A fuzzy bayesian network approach for risk analysis in process industries. *Process safety and environmental protection*, 111:507–519, 2017.

Sabine Zakel, Elisabeth Brandes, and Volkmar Schröder. Reliable safety characteristics of flammable gases and liquids–the database chemsafe. *Journal of Loss Prevention in the Process Industries*, 62:103914, 2019.

Appendix A: Delay time distributions

We consider a container filled with gaseous propane C_3H_8 in a closed room of constant volume initially at atmospheric pressure. We assume that the room is adiabatic, which means that all heat losses can be neglected.
Through a leak, a stoichiometric mixture of air-propane is formed in the room. Because of some incident, the ambient temperature jumps from 293.15 K (20 °C) to an initial temperature T_0 following a normal distribution: $T_0 \sim N(\mu_{T_0} = 524\ K, \sigma_{T_0} = 20\ K)$. We have $p(T_0 < 440K) < 1\text{E-}04$ and $p(T_0 > 600K) < 1\text{E-}04$. As a consequence, we can always consider that the initial temperature belongs to the interval $[T_{0,min} = 440\ \mathrm{K}; T_{0,max} = 600\ \mathrm{K}]$ in order to develop an approximation for the critical time.

We decided to describe the system through the theory of thermal explosion of Semenov [1]. According to it, the reaction progress leads to an increase in temperature which itself spawns an increase in the reaction rate that causes the temperature to rise ever faster (*thermal runaway*). Following this theory, the system can be described by the following equations:
$C_3H_8 + 5O_2 \rightarrow 3CO_2 + 4H_2O$, $\rho c_v \frac{\partial T}{\partial t} = (-\Delta u_m) r$ and
$r = -\frac{d[C_3H_8]}{dt} = -\frac{1}{5}\frac{d[O_2]}{dt} = Ae^{-\frac{Ea}{RT}}[C_3H_8]^a[O_2]^b$.
$[C_3H_8]$ and $[O_2]$ (mol/cm^3) are the concentrations of propane and oxygen, respectively, ρ (kg/cm^3) is the volumetric mass of the gaseous mixture, c_ν (kcal/(kg·K)) is its thermal capacity at constant volume, *T* (K) is the temperature, *t* (s) is the time, Δu_m (kcal/mol) is the molar reaction energy, *r* (mol/cm^3/s) is the reaction rate, *A* is the pre-exponential factor, *Ea* (kcal/mol) is the activation energy, *R* = 1.987E-03 kcal/(K mol) is the ideal gas constant, and *a* and *b* are reactant coefficients.

According to Westbrook and Dryer [2], the following parameter values can be employed:

Parameter	A_0 (mol, cm,s)	Ea_0 (kcal/mol)	*a*	*b*
Values	8.6E+11	30.0	0.1	1.65

They used laminar flame speeds at atmospheric pressures to calibrate them.

Further in the study, we shall assume that the parameters are uncertain and that $A \in [6.0\text{E+}11\ ;\ 8.0\text{E+}13]$ (mol, cm,s) and $Ea \in [27; 46]$ kcal/mol, for example because of analogies with other global reactions. We consider a stoichiometric mixture at atmospheric pressure (*p* = 1 atm = 101325 Pa), which means we have the following mole fractions:
$X_{C_3H_8} = \frac{[C_3H_8]}{[All\ species]} \approx 0.0406$, $X_{O_2} \approx 0.2029$, $X_{N_2} \approx 0.7565$.

We are interested in how quickly the mixture reaches the critical temperature[1] *Tc* = 766 K beyond which it would be impossible for someone to intervene to stop the explosion [3]. Let *tc* be the critical delay time defined as $T(tc) = T_c$.
$tc = tc_{A,Ea}(T_0)$ is a function of the kinetic parameters *A* and *Ea* and of the initial temperature T_0. For several values of the kinetic parameters *A* and *Ea*, we computed $tc_{A,Ea}(T_{0,min})$, $tc_{A,Ea}(T_{0,max})$ and $tc_{A,Ea}(T_{0,rand})$ for 40 random values of the initial temperature $T_{0,rand}$ uniformly chosen in the interval $[T_{0,min}; T_{0,max}]$ with the chemical kinetic software **Cantera [4]**. The results are shown in Figure 1.

1 also called ignition temperature in a technical context.

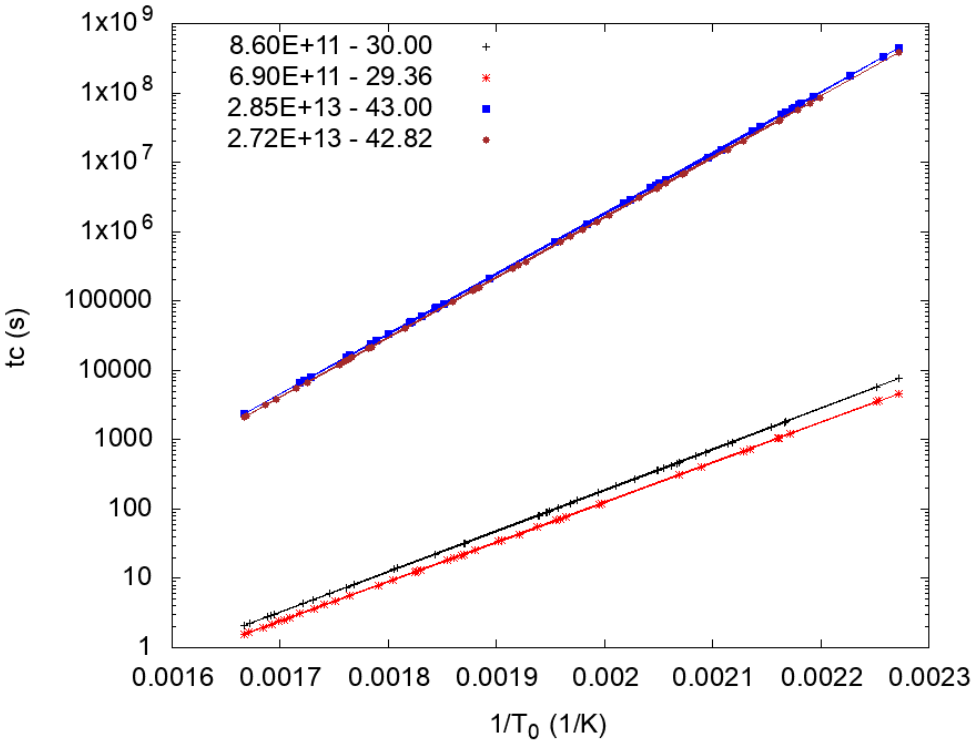


Figure 1: tc as a function of (1/T0) for different (A, Ea) shown in the legend.

We can see that the logarithmic values of the delay time can be well approximated by a linear function of *1/T₀*:

$log10(tc_{A,Ea}(T_0)) = a_{A,Ea}\frac{1}{T_0} + b_{A,Ea}$ with the coefficients

$a_{A,Ea} = \frac{log10\left(tc_{A,Ea}(T_{0,min})\right) - log10\left(tc_{A,Ea}(T_{0,max})\right)}{1/T_{0,min} - 1/T_{0,max}}$ and

$b_{A,Ea} = log10\left(tc_{A,Ea}(T_{0,min})\right) - a_{A,Ea}\frac{1}{T_{0,min}}$.

$log10\left(tc_{A,Ea}(T_{0,min})\right)$ and $log10\left(tc_{A,Ea}(T_{0,max})\right)$ also turn out to be well approximated by a bilinear function of *log10(A)* and *Ea*, as can be seen in Figure 2 and 3. They were obtained with a regular grid containing 30*30 values of (A, Ea). We used the results to create a **piece-wise** bilinear interpolation model of $log10\left(tc_{A,Ea}(T_{0,min})\right)$ and of $log10\left(tc_{A,Ea}(T_{0,max})\right)$ as a function of *A* and *Ea*.

Let us suppose we want to approximate $log10\left(tc_{A,Ea}(T_0)\right)$ for arbitrary values of *A, Ea* and T_0 belonging to the intervals defined above.
We start by predicting $log10\left(tc_{A,Ea}(T_{0,min})\right)$ and $log10\left(tc_{A,Ea}(T_{0,max})\right)$ with the piece-wise bilinear interpolation model.

We then compute the coefficients $a_{A,Ea} = \frac{log10\left(tc_{A,Ea}(T_{0,min})\right) - log10\left(tc_{A,Ea}(T_{0,max})\right)}{1/T_{0,min} - 1/T_{0,max}}$ and $b_{A,Ea} = log10\left(tc_{A,Ea}(T_{0,min})\right) - a_{A,Ea}\frac{1}{T_{0,min}}$.

We finally have: $log10\left(tc_{A,Ea}(T_0)\right) \approx a_{A,Ea}\frac{1}{T_0} + b_{A,Ea}$.

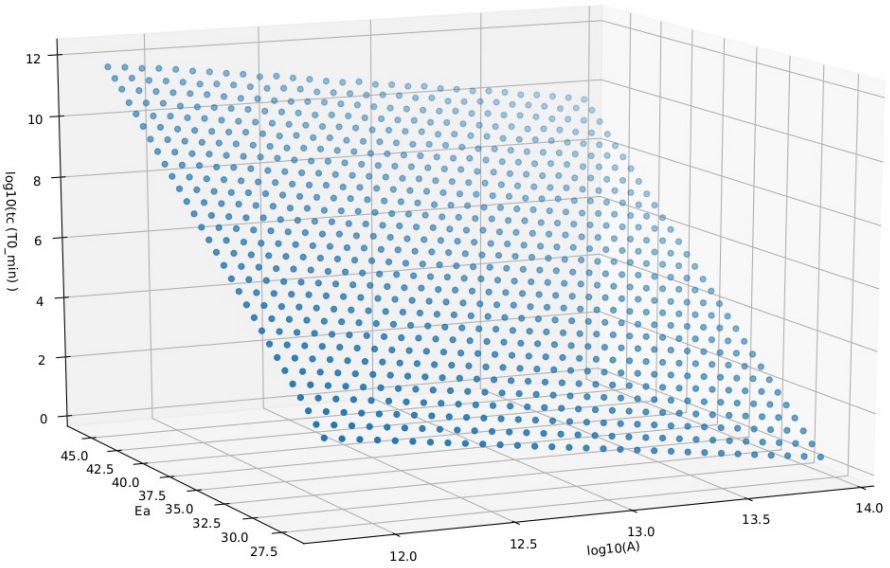


Figure 2: $log10\left(tc(T_{0,min})\right)$ as a function of (A, Ea)

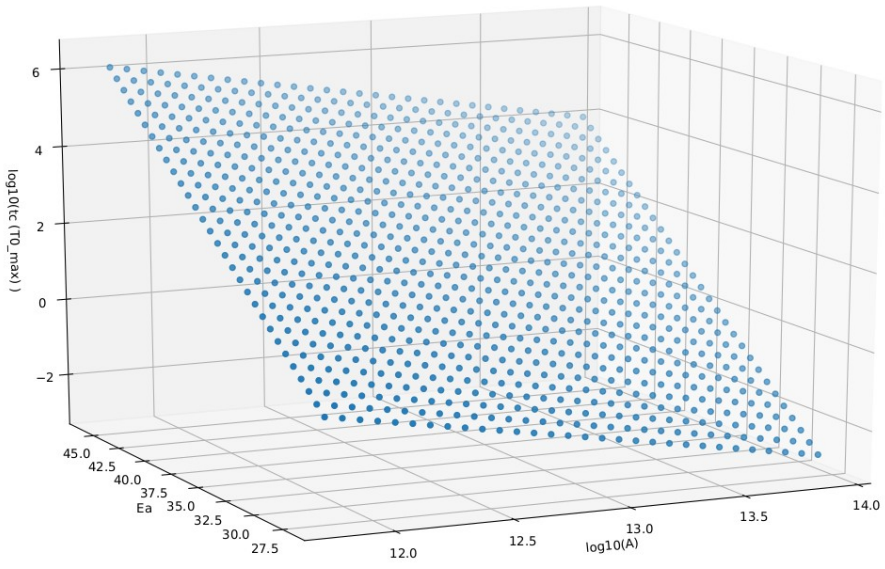


Figure 3: $log10\left(tc(T_{0,max})\right)$ as a function of (A, Ea)

The quality of the model has been tested by generating 3000 random values of (*A*, *Ea*) and T_0 with $A \sim U(A_{min}, A_{max})$, $Ea \sim U(Ea_{min}, Ea_{max})$ and $T_0 \sim N(\mu_{T_0}, \sigma_{T_0})$ whereby *U* designates a uniform probability distribution.

We systematically computed $tc_{A,Ea}(T_0)$ with Cantera and $tc_{pred,A,Ea}(T_0)$ with the two linear interpolation models and the relative difference $r = \frac{tc_{A,Ea}(T_0) - tc_{pred,A,Ea}(T_0)}{min\left(tc_{A,Ea}(T_0), tc_{pred,A,Ea}(T_0)\right)}$.

For more than 99.3% of the points, we have $|r| \leq 5\%$.

We have *max(r)* = 12.32% which is reached for a very high delay time superior to 2E+08 s that would be completely unproblematic. Given the fact that we are not interested in numerical accuracy but in understanding the behaviour of probabilistic approaches to thermal runaway hazard, we deemed that level of error to be acceptable.

We now want to compute the probability density distribution of *tc* for the true parameter values A_0 = 8.60E+11 and Ea_0 = 30.

We first compute the coefficients $a_{A_0,Ea_0} = \frac{log10\left(tc_{A_0,Ea_0}(T_{0,min})\right) - log10\left(tc_{A_0,Ea_0}(T_{0,max})\right)}{1/T_{0,min} - 1/T_{0,max}}$ and
$b_{A_0,Ea_0} = log10(tc_{A_0,Ea_0}(T_{0,min})) - a_{A_0,Ea_0}\frac{1}{T_{0,min}}$.

We then have $log10(tc_{A_0,Ea_0}(T_0)) = a_{A_0,Ea_0}\frac{1}{T_0} + b_{A_0,Ea_0}$ with
$T_0 \sim N(\mu_{T_0} = 525K, \sigma_{T_0} = 20K)$.

The cumulative probability distribution of t_c is given by

$$F(t) = p(tc \leq t) = p\big(log10(tc) \leq log10(t)\big) = p\Big(a_{A_0,Ea_0}\frac{1}{T_0} + b_{A_0,Ea_0} \leq log10(t)\Big)$$

$$F(t) = p(tc \leq t) = p\Big(\frac{1}{T_0} \leq \frac{log10(t) - b_{A_0,Ea_0}}{a_{A_0,Ea_0}}\Big) = p\Big(T_0 \geq \frac{a_{A_0,Ea_0}}{log10(t) - b_{A_0,Ea_0}}\Big)$$

$$F(t) = p(tc \leq t) = 1 - p\Big(T_0 \leq \frac{a_{A_0,Ea_0}}{log10(t) - b_{A_0,Ea_0}}\Big) = 1 - \Phi\Big(\frac{a_{A_0,Ea_0}}{log10(t) - b_{A_0,Ea_0}}, \mu_{T_0}, \sigma_{T_0}\Big)$$

where $\Phi(x, \mu_{T_0}, \sigma_{T_0})$ is the cumulative probability distribution of the normal distribution.

Hence $f(t) = F'(t) = \left(-\Phi\left(\frac{a_{A_0,Ea_0}}{log10(t) - b_{A_0,Ea_0}}, \mu_{T_0}, \sigma_{T_0}\right)\right)'$

$$f(t) = -\Big(\frac{a_{A_0,Ea_0}}{log10(t) - b_{A_0,Ea_0}}\Big)'\phi\Big(\frac{a_{A_0,Ea_0}}{log10(t) - b_{A_0,Ea_0}}, \mu_{T_0}, \sigma_{T_0}\Big)$$

$$f(t) = \frac{a_{A_0,Ea_0}ln(10)}{t\big(b_{A_0,Ea_0}ln(10) - ln(t)\big)^2}\phi\Big(\frac{a_{A_0,Ea_0}}{log10(t) - b_{A_0,Ea_0}}, \mu_{T_0}, \sigma_{T_0}\Big)$$

We can also compute the quantiles tc_{25}, tc_{50} and tc_{75} (with $F(tc_{25}) = 0.25$, $F(tc_{50}) = 0.50$ and $F(tc_{75}) = 0.75$) and $p_c = p(tc < 30s) = F(30)$ which is the probability that the thermal runaway delay would be so small that it would be very hard for someone to step in.

We have $0.25 = p(tc \leq tc_{25}) = 1 - \Phi\Big(\frac{a_{A_0,Ea_0}}{log10(tc_{25}) - b_{A_0,Ea_0}}, \mu_{T_0}, \sigma_{T_0}\Big)$

$$\Phi\left(\frac{a_{A_0,Ea_0}}{log10(tc_{25}) - b_{A_0,Ea_0}}, \mu_{T_0}, \sigma_{T_0}\right) = 0.75.$$

We thus have $\frac{a_{A_0,Ea_0}}{log10(tc_{25}) - b_{A_0,Ea_0}} = Z_{0.75}(\mu_{T_0}, \sigma_{T_0})$, where $Z_{0.75}(\mu_{T_0}, \sigma_{T_0})$ is a quantile of $N(\mu_{T_0}, \sigma_{T_0})$.

This leads to $tc_{25} = 10^{\frac{a_{A_0,Ea_0}}{Z_{0.75}(\mu_{T_0},\sigma_{T_0})} + b_{A_0,Ea_0}}$. Likewise, we have

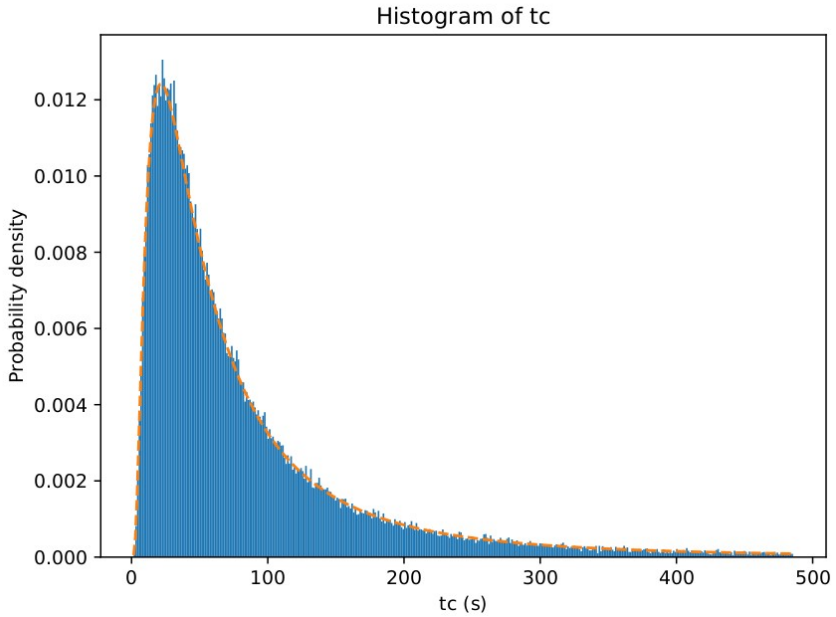
$tc_{75} = 10^{\frac{a_{A_0,Ea_0}}{Z_{0.25}(\mu_{T_0},\sigma_{T_0})} + b_{A_0,Ea_0}}$ and $tc_{50} = 10^{\frac{a_{A_0,Ea_0}}{Z_{0.50}(\mu_{T_0},\sigma_{T_0})} + b_{A_0,Ea_0}}$.

Of course, these formula are also valid for any $A \in [A_{min}, A_{max}]$ and $Ea \in [Ea_{min}, Ea_{max}]$.

We compared the analytical functions with an histogram obtained by generating 10,000 values of $T_0 \sim N(\mu_{T_0} = 525K, \sigma_{T_0} = 20K)$ and computing the corresponding *tc* for (A, Ea) via the two linear interpolations. We obtained the following results:

***A* = 8.6E+11 (mol, cm, s) - *Ea* = 30 kcal/mol**	tc_{25}	tc_{50}	tc_{75}	$p_{critical}$
Analytical	28.84	55.04	108.68	0.2636
Numerical	28.94	54.94	107.72	0.2625

Figure 4:



Probability distribution for A_0 = **8.6E+11** *and Ea_0 =* **30**

We also computed the probability distribution for two other (*A*, *Ea*) chosen randomly.

***A* = 6.14E+12 - *Ea* = 37.36 kcal/mol**	tc_{25}	tc_{50}	tc_{75}	$p_{critical}$
Analytical	3109.14	7071.60	16797.92	7.75E-08
Numerical	3099.12	7069.12	16815.32	0

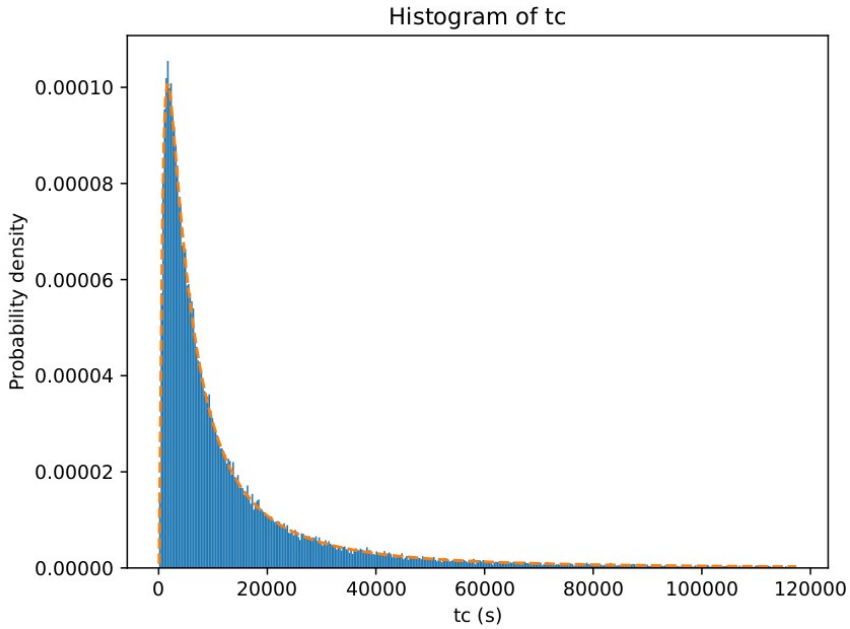
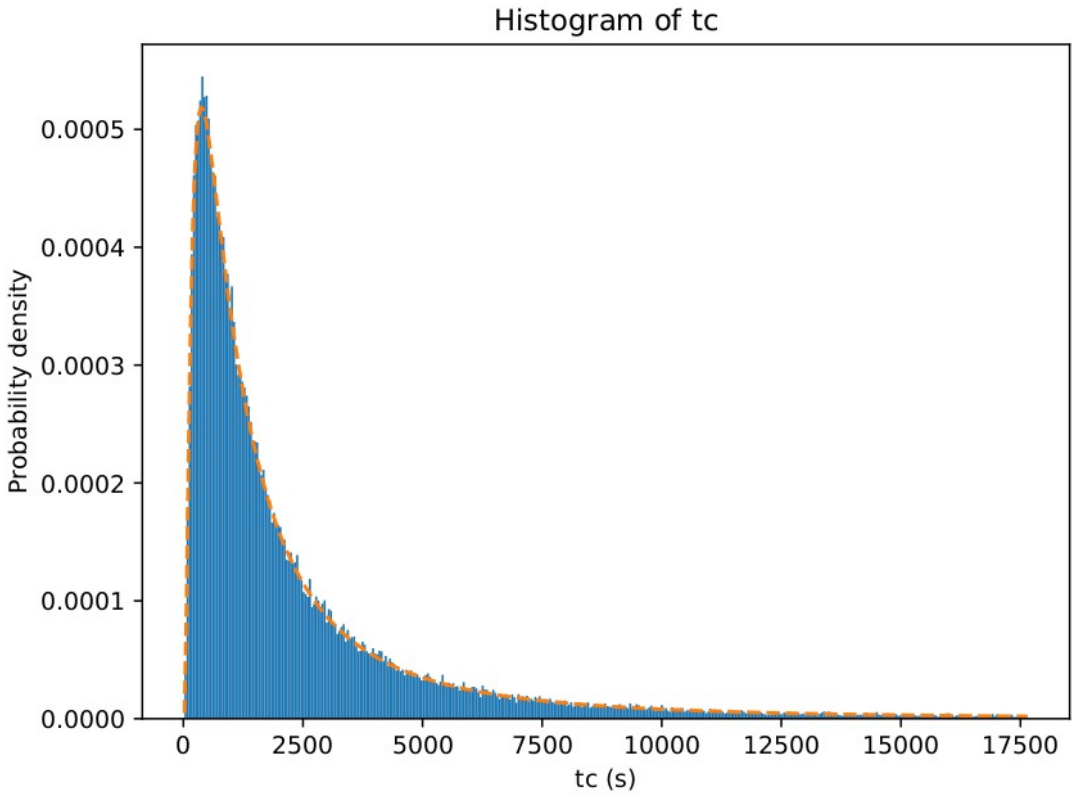


Figure 5: Probability distribution for A = **6.14E+12** *and Ea =* **37.36**

***A* = 3.06E+12 - *Ea* = 34.82 kcal/mol**	tc_{25}	tc_{50}	tc_{75}	$p_{critical}$
Analytical	625.17	1336.97	2976.44	8.66E-05
Numerical	625.34	1340.69	2990.27	9.00E-05

Figure 6:



Probability distribution for A = **3.06E+12** *and Ea =* **34.82**

***A* = 2.99E+13 - *Ea* = 43.18 kcal/mol**	tc_{25}	tc_{50}	tc_{75}	$p_{critical}$
Analytical	1.24E+05	3.24E+05	8.86E+05	0
Numerical	1.24E+05	3.20E+05	8.84E+05	0

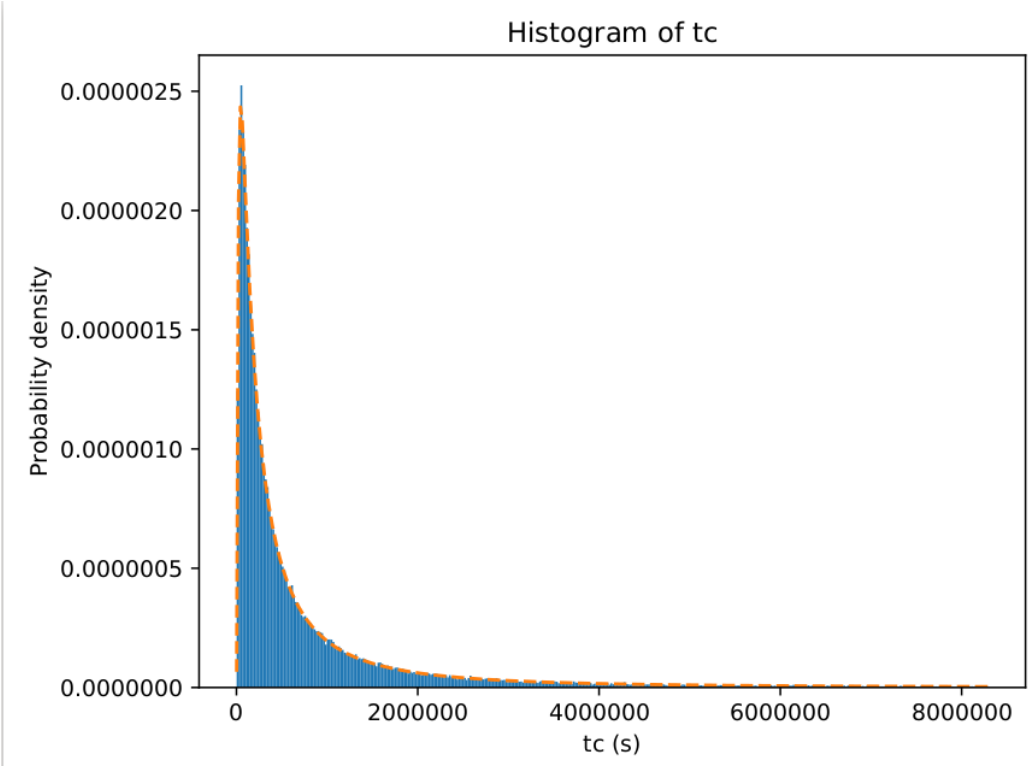


Figure 7: Probability distribution for A = **2.99E+13** *and Ea =* **43.18**

Tests performed with other parameter values led to an equally satisfying agreement, thereby validating the analytical formula.

[1] N.N. Semenov, Thermal theory of combustion and explosion, Natl. Advis. Comm. Aeronaut. -- Tech. Memo. (1942) 52-.

[2] C.K. Westbrook, F.L. Dryer, Simplified reaction mechanisms for the oxidation of hydrocarbon fuels in flames, Combust. Sci. Technol. 27 (1981) 31–43.

[3] R.J. Reed, North American Combustion Handbook, vol. 1: Combustion, Fuels, Stoichiometry, Heat Transf. Fluid Flow. (1986).

[4] D.G. Goodwin, Cantera C++ user's guide, Calif. Inst. Technol. (2002).

Appendix B: priors

In order to perform the imprecise Bayesian analysis, we decided to use sixth priors which are uniform with respect to the following variables:
- first prior: *log10(A) - Ea*
- second prior: *A – Ea*
- third prior: *1/log10(A) - Ea*
- fourth prior: *1/log10(A) – 1/Ea*
- fifth prior: *1/A – Ea*
- sixth prior: *A – 1/Ea*

Let us pose $x = log10(A)$. Let us first consider the univariate priors $f_{log10(A)}$, $f_{\frac{1}{log10(A)}}$, f_A, $f_{\frac{1}{A}}$, f_{Ea}, $f_{\frac{1}{Ea}}$ which are uniform with respect to the indexed variables. We shall express the corresponding cumulative distribution functions and then the pdf with respect to *x* and *Ea*.

We straightforwardly have $f_{log10(A)}(log10(A)) = \frac{1}{log10(A_{max}) - log10(A_{min})}$ and $f_{Ea}(Ea) = \frac{1}{Ea_{max} - Ea_{min}}$.

For $f_{\frac{1}{log10(A)}}$**,** we have $F_{\frac{1}{log10(A)}}(x) = p\big(log10(A) \le x\big) = p\big(\frac{1}{log10(A)} \ge \frac{1}{x}\big) = 1 - F_{\frac{1}{log10(A)}}\big(\frac{1}{x}\big)$

$$F_{\frac{1}{log10(A)}}(x) = 1 - \frac{1}{\frac{1}{log10(A_{min})} - \frac{1}{log10(A_{max})}}\big(\frac{1}{x} - \frac{1}{log10(A_{max})}\big)$$

$$f_{\frac{1}{log10(A)}}(x) = \frac{1}{\frac{1}{log10(A_{min})} - \frac{1}{log10(A_{max})}}\frac{1}{x^2}$$

$$\rightarrow f_{\frac{1}{log10(A)}}\big(log10(A)\big) = \frac{1}{\frac{1}{log10(A_{min})} - \frac{1}{log10(A_{max})}}\frac{1}{log10(A)^2}$$

For f_A, we have $F_A(x) = \frac{10^x - A_{min}}{A_{max} - A_{min}} \rightarrow f_A(x) = \frac{ln(10)10^x}{A_{max} - A_{min}}$,

$f_A\big(log10(A)\big) = \frac{ln(10)10^{log10(A)}}{A_{max} - A_{min}}$.

For $f_{\frac{1}{A}}$, we have

$$F_{\frac{1}{A}}(x) = \mathrm{p}\ \big(log10(A) \le x\big) = p\big(A \le 10^x\big) = \mathrm{p}\ \big(\tfrac{1}{A} \ge \tfrac{1}{10^x}\big) = 1 - F_{\frac{1}{A}}\big(\tfrac{1}{10^x}\big)$$

$$\mathrm{F}_{\frac{1}{A}}(x) = 1 - \frac{\frac{1}{10^x} - \frac{1}{A_{max}}}{\frac{1}{A_{min}} - \frac{1}{A_{max}}} \rightarrow \mathrm{f}_{\frac{1}{A}}(x) = \frac{ln(10)10^{-x}}{\frac{1}{A_{min}} - \frac{1}{A_{max}}} \rightarrow \mathrm{f}_{\frac{1}{A}}\big(log10(A)\big) = \frac{ln(10)10^{-log10(A)}}{\frac{1}{A_{min}} - \frac{1}{A_{max}}}$$

$$F_{\frac{1}{Ea}}(Ea) = 1 - \frac{\frac{1}{Ea} - \frac{1}{Ea_{max}}}{\frac{1}{Ea_{min}} - \frac{1}{Ea_{max}}} \rightarrow f_{\frac{1}{Ea}}(Ea) = \frac{1}{Ea^2}\frac{1}{\frac{1}{Ea_{min}} - \frac{1}{Ea_{max}}}.$$

For the first prior (uniform with respect to *log10(A)* and *Ea)*, we have

$$f_{0,1}(log10(A), Ea) = f_{log10(A),Ea}(log10(A), Ea) = f_{log10(A)}\big(log10(A)\big) f_{Ea}(Ea)$$

$$f_{0,1}(log10(A), Ea) = \frac{1}{log10(A_{max}) - log10(A_{min})}\frac{1}{Ea_{max} - Ea_{min}}$$

For the second prior (uniform with respect to A and *Ea)*, we have

$$f_{0,2}(log10(A), Ea) = f_{A,Ea}(log10(A), Ea) = f_{A}\big(log10(A)\big) f_{Ea}(Ea)$$

$$f_{0,2}(log10(A), Ea) = \frac{ln(10)}{A_{max} - A_{min}}\frac{1}{Ea_{max} - Ea_{min}} 10^{log10(A)}$$

For the third prior (uniform with respect to *1/log10(A)* and *Ea)*, we have

$$f_{0,3}(log10(A), Ea) = f_{\frac{1}{log10(A)},Ea}(log10(A), Ea) = f_{\frac{1}{log10(A)}}\big(log10(A)\big) f_{Ea}(Ea)$$

$$f_{0,3}(log10(A), Ea) = \frac{1}{\frac{1}{log10(A_{min})} - \frac{1}{log10(A_{max})}}\frac{1}{Ea_{max} - Ea_{min}}\frac{1}{log10(A)^2}$$

For the fourth prior (uniform with respect to *1/log10(A)* and *1/Ea)*, we have

$$f_{0,4}(log10(A), Ea) = f_{\frac{1}{log10(A)},\frac{1}{Ea}}(log10(A), Ea) = f_{\frac{1}{log10(A)}}\big(log10(A)\big) f_{\frac{1}{Ea}}(Ea)$$

$$f_{0,4}(log10(A), Ea) = \frac{1}{\frac{1}{log10(A_{min})} - \frac{1}{log10(A_{max})}}\frac{1}{\frac{1}{Ea_{min}} - \frac{1}{Ea_{max}}}\frac{1}{log10(A)^2}\frac{1}{Ea^2}$$

For the fifth prior (uniform with respect to *1/A* and *Ea)*, we have

$$f_{0,5}(log10(A), Ea) = f_{\frac{1}{A},Ea}(log10(A), Ea) = f_{\frac{1}{A}}\big(log10(A)\big) f_{Ea}(Ea)$$

$$f_{0,5}(log10(A), Ea) = \frac{ln(10)}{\frac{1}{A_{min}} - \frac{1}{A_{max}}}\frac{1}{Ea_{max} - Ea_{min}} 10^{-log10(A)}$$

For the sixth prior (uniform with respect to *A* and *1/Ea)*, we have

$$f_{0,6}(log10(A), Ea) = f_{A,\frac{1}{Ea}}(log10(A), Ea) = f_{A}\big(log10(A)\big) f_{\frac{1}{Ea}}(Ea)$$

$$f_{0,6}(log10(A), Ea) = \frac{ln(10)}{A_{max} - A_{min}}\frac{1}{\frac{1}{Ea_{min}} - \frac{1}{Ea_{max}}} 10^{log10(A)}\frac{1}{Ea^2}$$

The six priors are shown in the following figures.

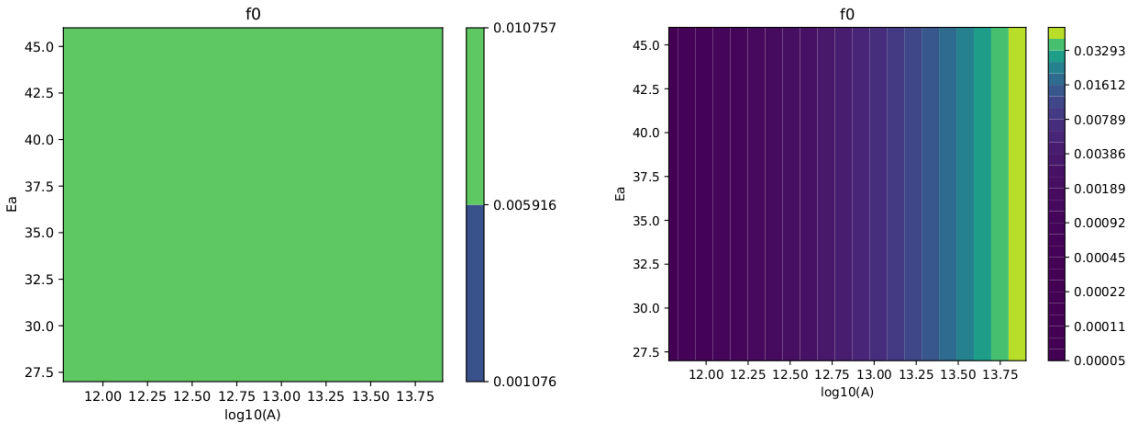


Figure 1: $f_{0,1}(log10(A), Ea) = f_{log10(A),Ea}(log10(A), Ea)$ - $f_{0,2}(log10(A), Ea) = f_{A,Ea}(log10(A), Ea)$

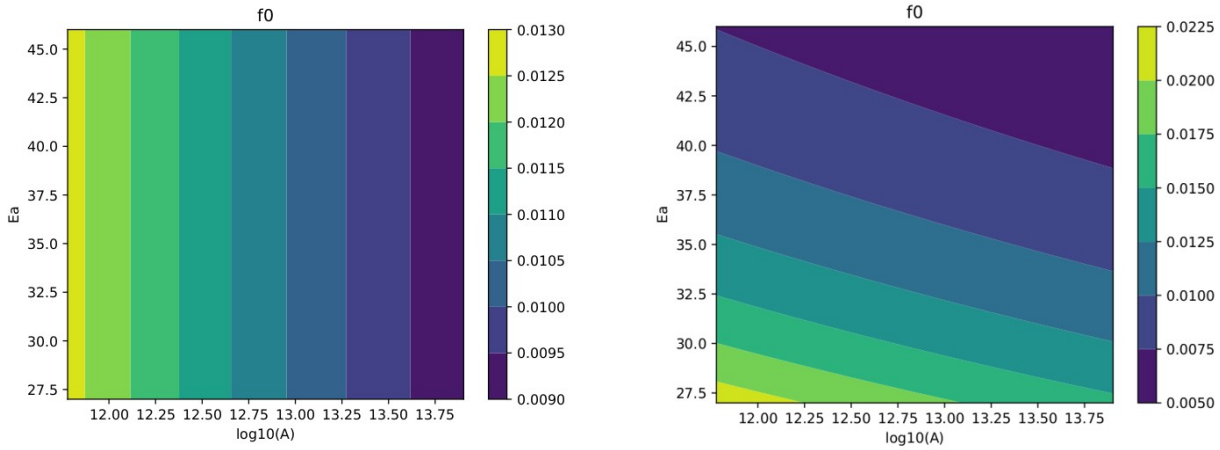


Figure 2: $f_{0,3}(log10(A), Ea) = f_{\frac{1}{log10(A)},Ea}(log10(A), Ea)$ - $f_{0,4}(log10(A), Ea) = f_{\frac{1}{log10(A)},\frac{1}{Ea}}(log10(A), Ea)$

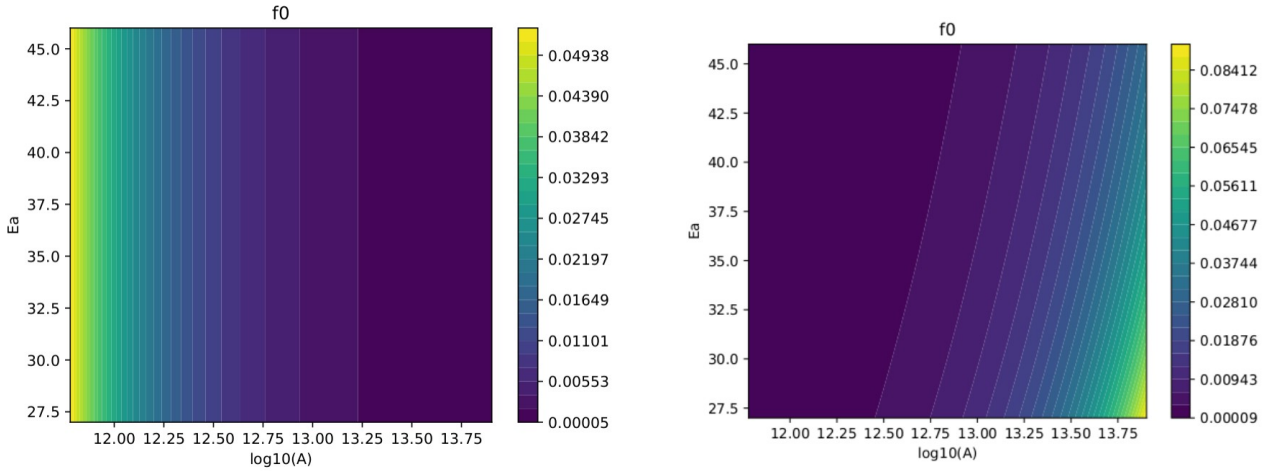


Figure 3: $f_{0,5}(log10(A), Ea) = f_{\frac{1}{A},Ea}(log10(A), Ea)$ - $f_{0,6}(log10(A), Ea) = f_{A,\frac{1}{Ea}}(log10(A), Ea)$

The integrals of the priors were computed over $A \in [6.0\text{E}{+11}\ ;\ 8.0\text{E}{+13}]$ mol/m^3/s and $Ea \in [27; 46]$.

Prior	$f_{0,1}$	$f_{0,2}$	$f_{0,3}$	$f_{0,4}$	$f_{0,5}$	$f_{0,6}$
Integral	1.00100075	1.00100422	1.00100064	1.00114586	1.00100054	1.00114796

Appendix C: virtual experimental data

We created “experimental” data allowing us to determine posterior probability distributions of *A* and *Ea*.
We considered a constant-volume reactor at atmospheric pressure with a very diluted mixture of propane and oxygen: X_{C3H8} = 1E-05, X_{O2} = 5E-05 and X_{N2} = 0.99994.
Under such a high dilution, the temperature remains nearly constant so that analytical expressions of the profile of $X_{C3H8}(t)$ are available. It has been rigorously verified that they are virtually identical to the numerical results of Cantera.
Using the “true” values A_0 = 8.60E+11 and Ea_0 = 30.00 kcal/mol, we generated mole fraction profiles of C_3H_8 at four different temperatures in the range [1135; 2249] K. We used the model described in subsection 2.1 of the article and in appendix A.
We then randomly chose several time points t_i and generated normally distributed noise in such a way that $X_{C_3H_8,exp}(t_i) = X_{C_3H_8}(t_i) + \epsilon_i$ and $\epsilon_i \sim N\left(0, \sigma_r.X_{C3H8}(t_i)\right)$. σ_r is the relative standard deviation. The profile of propane $X_{C3H8}(t)$ has always been computed with A_0 and Ea_0.

We considered four situations:

A) we have no data whatsoever, we only know that $A \in [6.0\text{E}{+11};8.0\text{E}{+13}]$ mol/m^3/s and $Ea \in [27; 46]$ kcal/mol.
B) We have one profile of C_3H_8 with 6 time points measured at 1845 K with σ_r = 25%.
C) We have two profiles of C_3H_8 with 6 time points measured at 1135 K and 2249 K with σ_r = 25%.
D) We have four profiles of C_3H_8 with 10 time points measured at 1135 K, 1478 K, 1845 K, and 2249 K with σ_r = 6%.

The different profiles are shown in the figures below.

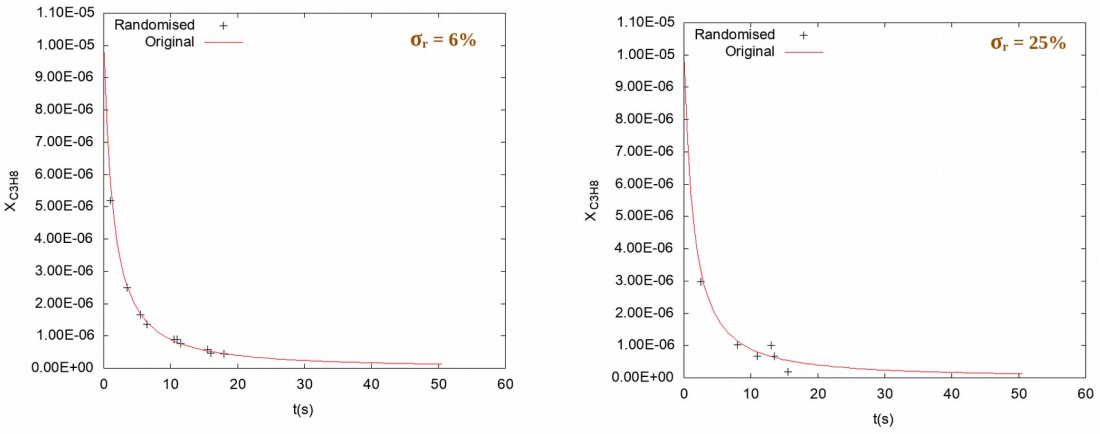


Figure 1: T_0 = 1135 K.

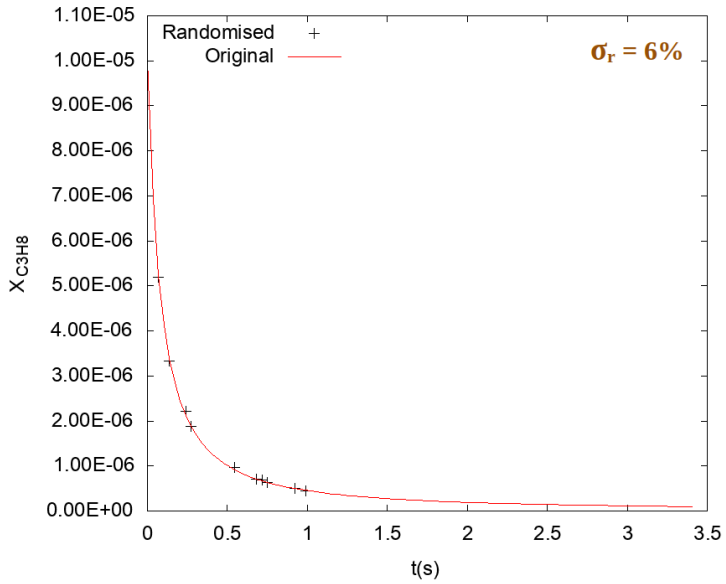
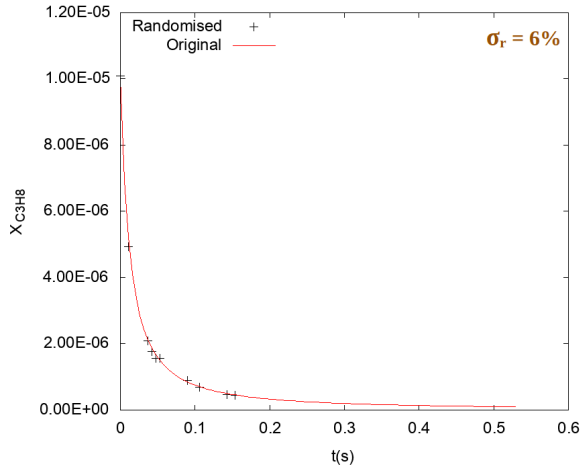


Figure 2: T_0 = 1478 K.



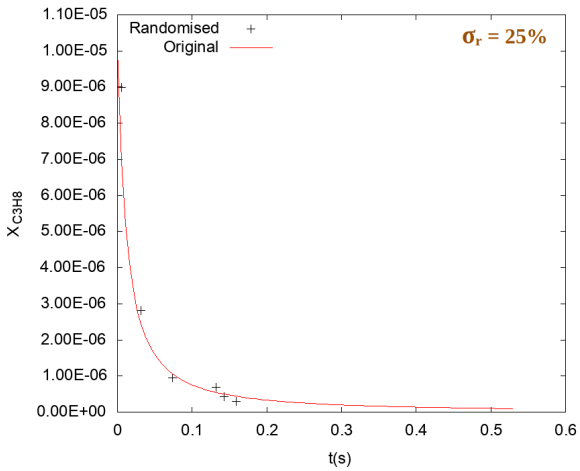
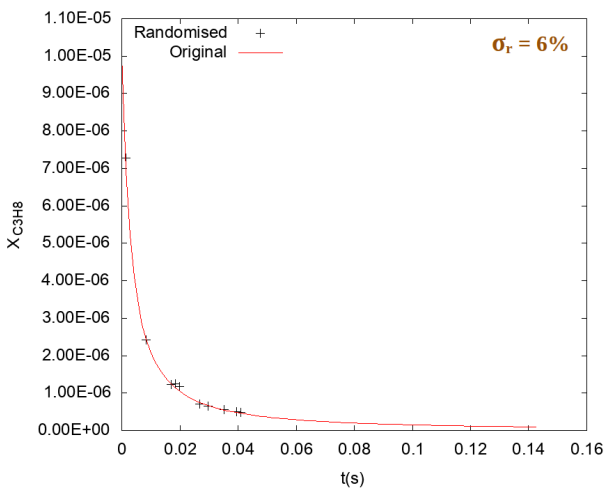


Figure 3: T_0 = 1845 K.



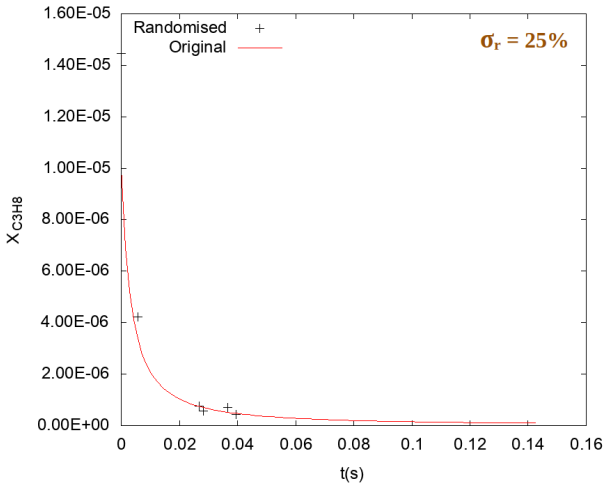


Figure 4: T_0 *=2249 K.*

The likelihood of the experimental data given the parameters is:

$$L(data|A,Ea) = \prod_{i=1}^{m}\prod_{j=1}^{n_{t,i}} \frac{1}{\sqrt{2\pi\sigma_{i,j}^2}} exp\left(-\frac{1}{2\sigma_{i,j}^2}\left(X_{C3H8,i}(t_j,A,Ea) - X_{C3H8,exp,i}(t_j)\right)^2\right)$$

where *m* is the number of experiments, $n_{t,i}$ is the number of time points for the *i*-th experiment and $\sigma_{i,j} = \sigma_{r,i} X_{C3H8,i}(t_j, A_0, Ea_0)$ is the local standard deviation.

The log-likelihood is given by

$$l(data|A,Ea) = \sum_{i=1}^{m}\sum_{j=1}^{n_{t,i}}\left(ln\left(\frac{1}{\sqrt{2\pi\sigma_{i,j}^2}}\right) - \frac{1}{2\sigma_{i,j}^2}\left(X_{C3H8,i}(t_j,A,Ea) - X_{C3H8,exp,i}(t_j)\right)^2\right)$$

$$l(data|A,Ea) = \sum_{i=1}^{m}\sum_{j=1}^{n_{t,i}}\left(-\frac{1}{2}ln\left(2\pi\sigma_{i,j}^2\right) - \frac{1}{2\sigma_{i,j}^2}\left(X_{C3H8,i}(t_j,A,Ea) - X_{C3H8,exp,i}(t_j)\right)^2\right)$$

The log-likelihood for situation B, C and D can be visualised in the figures just below. They were all obtained with **300*300** values of *A* and *Ea*.

We can see that the likelihood function becomes sharper and sharper as more precise measurements are used to compute it, which is completely expected.

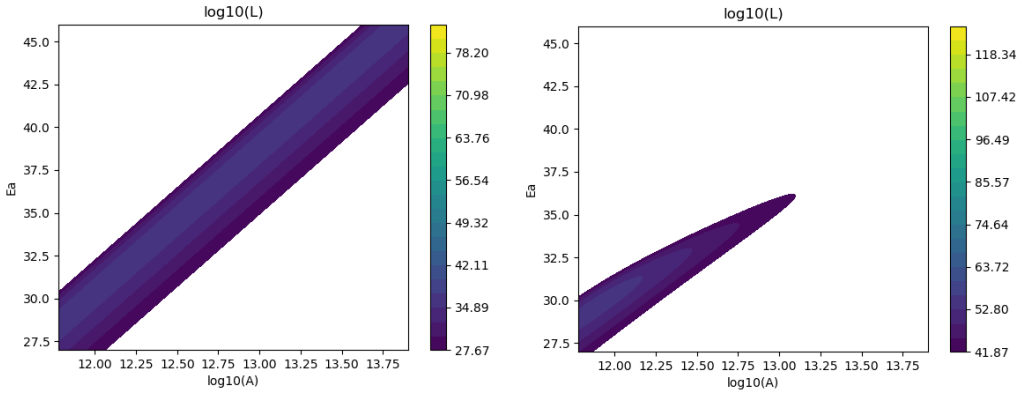


Figure 5: Situation B - Situation C

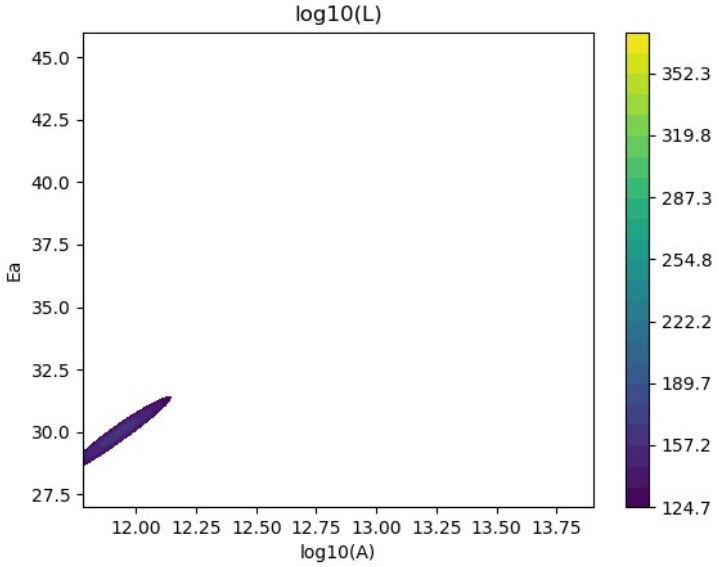


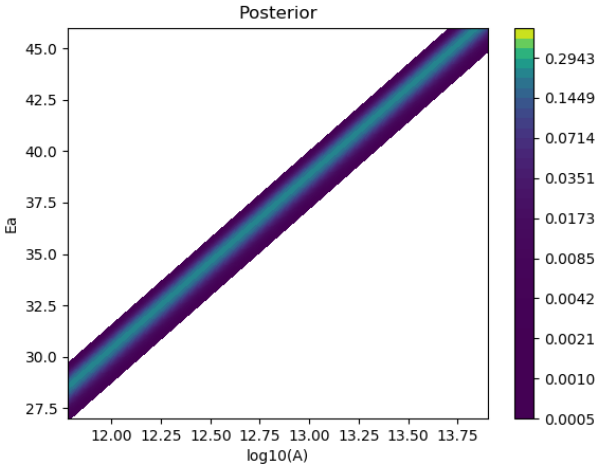
Figure 6: Situation D

For each situation B, C and D, we obtained 6 posteriors based on the six priors mentioned in Appendix B.

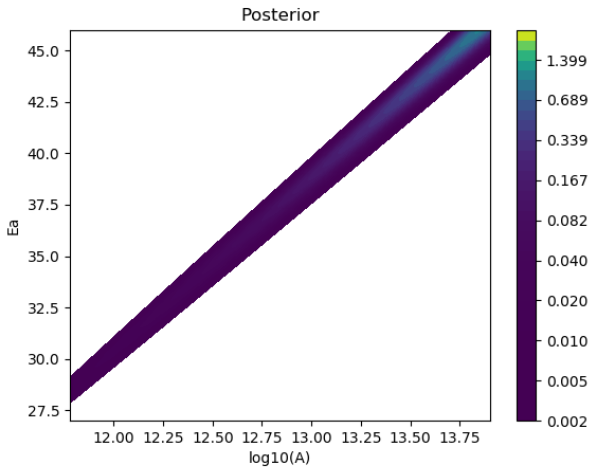
$f_{B,1}\left(log10(A), Ea\right) = \frac{L\left(B|log10(A), Ea\right) f_{0,1}\left(log10(A), Ea\right)}{\int_{log10(A)} \int_{Ea} L\left(B|log10(A), Ea\right) f_{0,1}\left(log10(A), Ea\right) dlog10(A) dEa}$ is the posterior based on the first prior $f_{0,1}$ and the experiment B.

$f_{C,4}\left(log10(A), Ea\right) = \frac{L\left(C|log10(A), Ea\right) f_{0,4}\left(log10(A), Ea\right)}{\int_{log10(A)} \int_{Ea} L\left(C|log10(A), Ea\right) f_{0,4}\left(log10(A), Ea\right) dlog10(A) dEa}$ is the posterior based on the fourth prior $f_{0,4}$ and the experiment C.

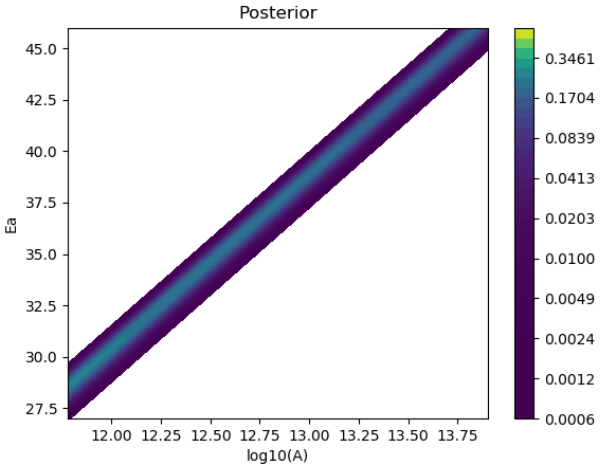
And so on.



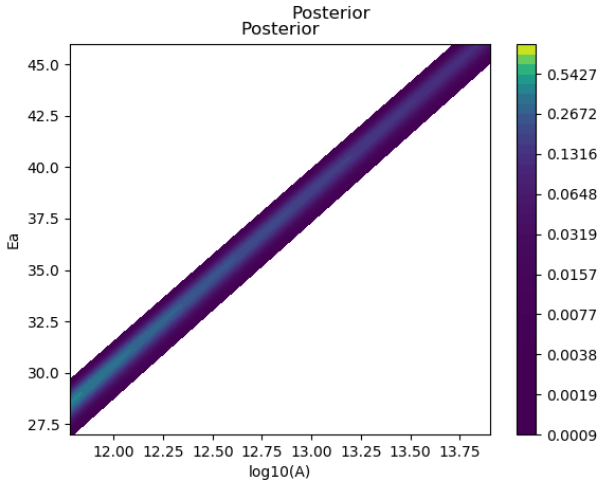
$\mathbf{f_{B,1}}$



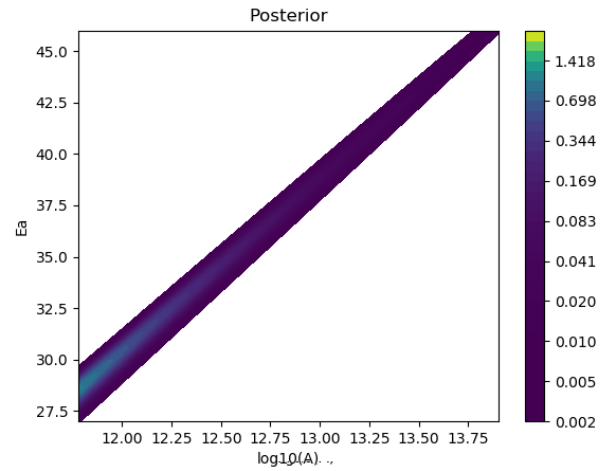
$\mathbf{f_{B,2}}$



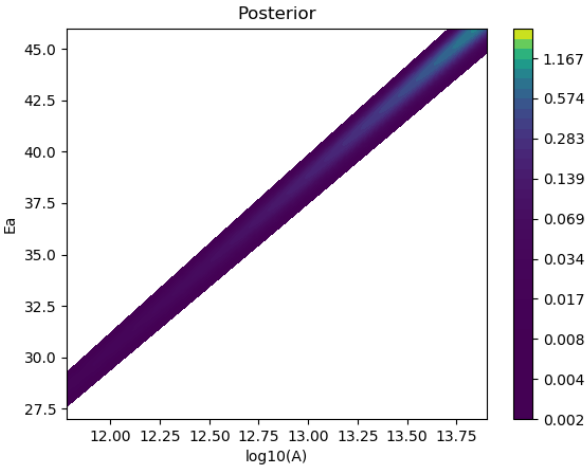
$\mathbf{f_{B,3}}$



$\mathbf{f_{B,4}}$

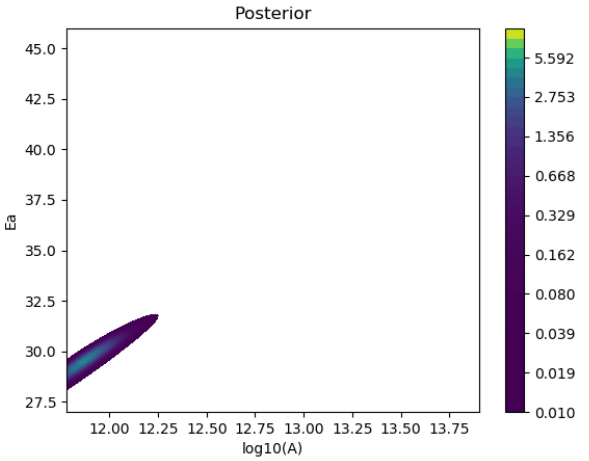


$\mathbf{f_{B,5}}$

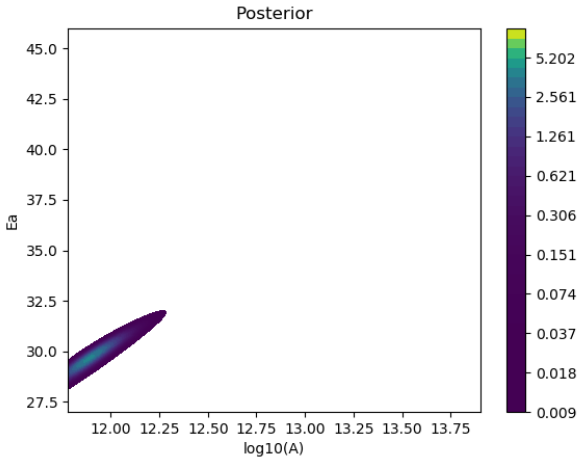


$\mathbf{f_{B,6}}$

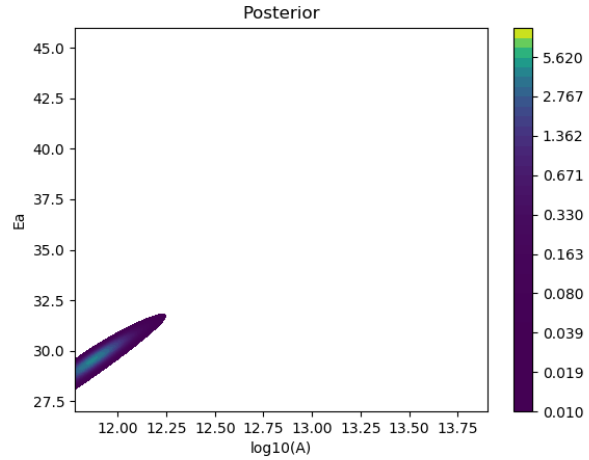
In situation B, *log10(A)* and *Ea* are strongly correlated in most posteriors and many values of (*log10(A)* , *Ea*) are equally compatible with the available evidence.



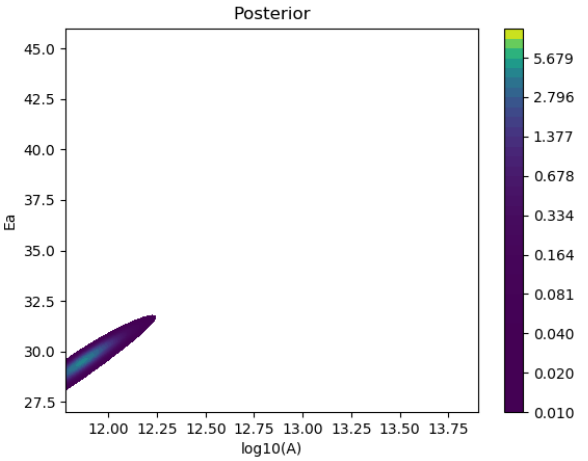
$\mathbf{f_{C,1}}$



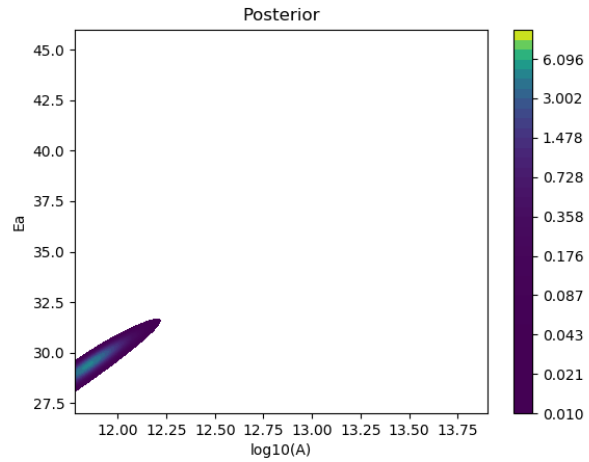
$\mathbf{f_{C,2}}$



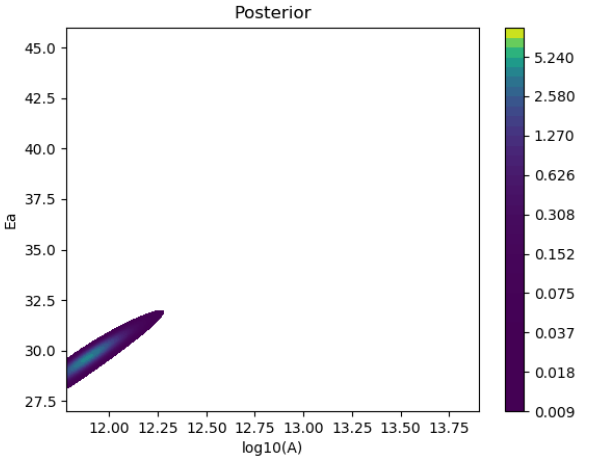
$\mathbf{f_{C,3}}$



$\mathbf{f_{C,4}}$

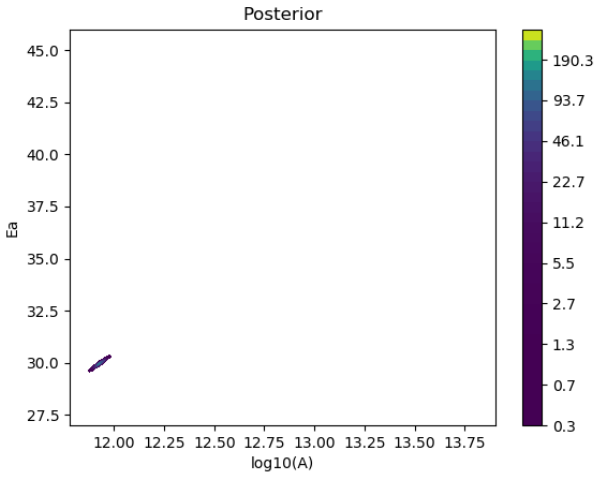


$\mathbf{f_{C,5}}$

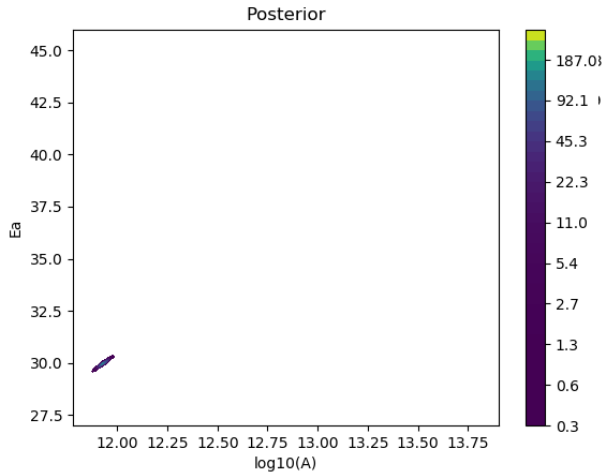


$\mathbf{f_{C,6}}$

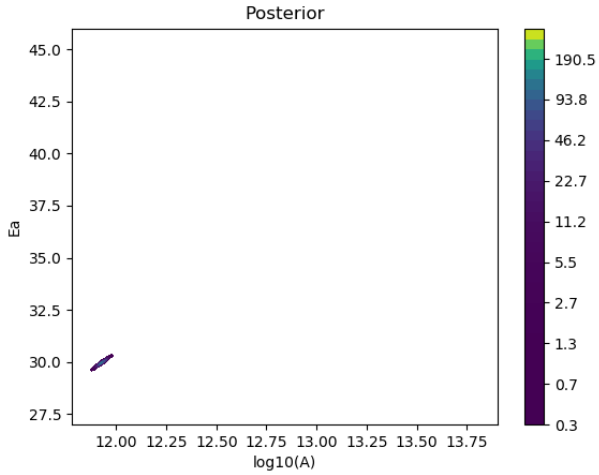
The range of probable parameter values is considerably smaller in situation C but this could still lead to inaccurate predictions.



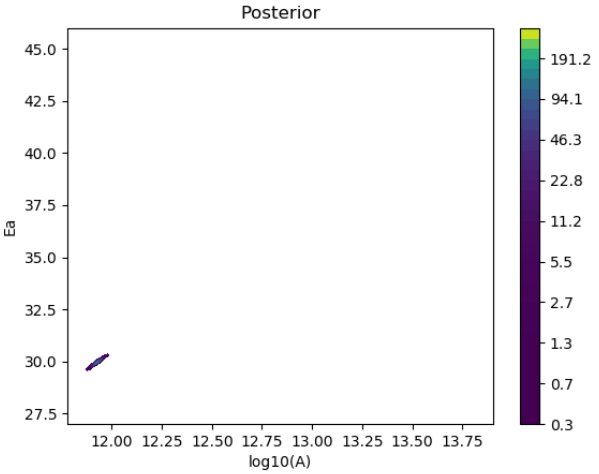
$\mathbf{f_{D,1}}$



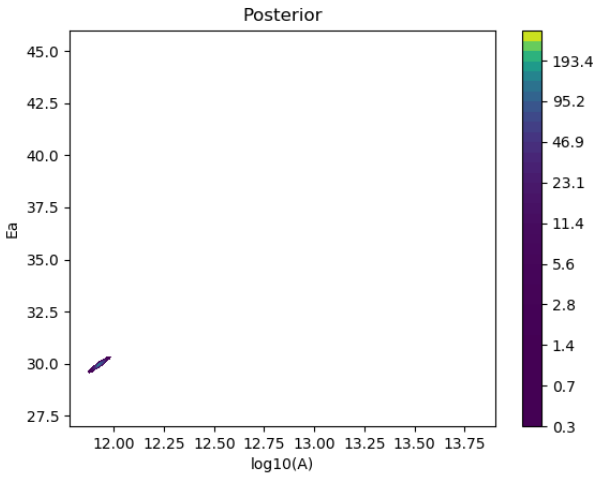
$\mathbf{f_{D,2}}$



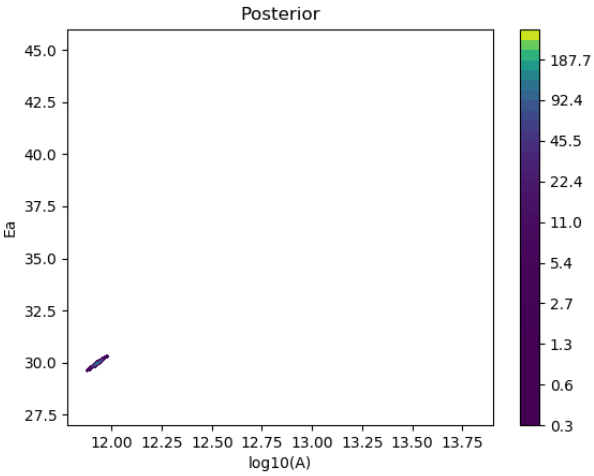
$\mathbf{f_{D,3}}$



$\mathbf{f_{D,4}}$



$\mathbf{f_{D,5}}$



$\mathbf{f_{D,6}}$

The range of probable parameter values is very small in situation D in all posteriors, which means that the parameters are known with a high degree of accuracy.

Appendix D: Bayesian estimation of the delay time distribution

We are interested in how quickly the mixture reaches the critical temperature *Tc* = 766 K beyond which it would be impossible for a technician to intervene to stop the explosion. *tc* is the critical delay time defined as $T(tc) = T_c$. As explained in appendix A, the initial temperature follows a normal distribution $T_0 \sim N(\mu_{T_0} = 524K, \sigma_{T_0} = 20K)$. We can limit ourselves to the interval $[T_{0,min} = 440 \text{ K}; T_{0,max} = 600 \text{ K}]$ as $p(T < T_{0,min} \cup T > T_{0,max})$ < 2E-04.
The true values of the kinetic parameters are A_0 = 8.6E+11 and Ea_0 = 30.0.

We suppose they are unknown to the safety engineer so that $A \in [6.0\text{E}{+}11 \ ; \ 8.0\text{E}{+}13]$ and $Ea \in [27; 46]$ kcal/mol. $tc_{A,Ea}(T_{0,min})$ and $tc_{A,Ea}(T_{0,max})$ can be predicted with a high degree of accuracy thanks to piecewise bilinear interpolation.
It is then possible to predict $tc_{A,Ea}(T_0)$ through a simple linear interpolation:

$log10(tc_{A,Ea}(T_0)) = a_{A,Ea}\frac{1}{T_0} + b_{A,Ea}$ with the coefficients

$a_{A,Ea} = \frac{log10\big(tc_{A,Ea}(T_{0,min})\big) - log10\big(tc_{A,Ea}(T_{0,max})\big)}{1/T_{0,min} - 1/T_{0,max}}$ and

$b_{A,Ea} = log10\big(tc_{A,Ea}(T_{0,min})\big) - a_{A,Ea}\frac{1}{T_{0,min}}$.

For *A* and *Ea*, the probability distribution of t_c is given by
$f(tc|A, Ea) = \frac{a_{A,Ea} ln(10)}{tc\big(b_{A,Ea} ln(10) - ln(tc)\big)^2} \phi\Big(\frac{a_{A,Ea}}{log10(tc) - b_{A,Ea}}, \mu_{T_0}, \sigma_{T_0}\Big)$.

We are interested in the quantiles tc_{25}, tc_{50} and tc_{75} which are given by the following formula $tc_{25} = 10^{\frac{a_{A,Ea}}{Z_{0.75}(\mu_{T_0}, \sigma_{T_0})} + b_{A,Ea}}$, $tc_{50} = 10^{\frac{a_{A,Ea}}{Z_{0.50}(\mu_{T_0}, \sigma_{T_0})} + b_{A,Ea}}$ and $tc_{75} = 10^{\frac{a_{A,Ea}}{Z_{0.25}(\mu_{T_0}, \sigma_{T_0})} + b_{A,Ea}}$

We also want to know $p_c = p(tc < 30s) = F(30)$ which is the probability that the thermal runaway delay would be so small that it would be very hard for a technician to step in.
F is the cumulative probability distribution of t_c given by

$F(tc|A, Ea) = 1 - \Phi\Big(\frac{a_{A,Ea}}{log10(tc) - b_{A,Ea}}, \mu_{T_0}, \sigma_{T_0}\Big)$.

For a given (*A*, *Ea*), we can determine $tc_{0.001}$ and $tc_{0.999}$ such that $F(tc_{0.001}|A, Ea) = 0.001$ and $F(tc_{0.999}|A, Ea) = 0.999$.
We have $0.001 = p(tc \leq tc_{0.001}) = 1 - \Phi\Big(\frac{a_{A,Ea}}{log10(tc_{0.001}) - b_{A,Ea}}, \mu_{T_0}, \sigma_{T_0}\Big)$

$\Phi\Big(\frac{a_{A,Ea}}{log10(tc_{0.001}) - b_{A,Ea}}, \mu_{T_0}, \sigma_{T_0}\Big) = 0.999$.

We thus have $\frac{a_{A,Ea}}{log10(tc_{0.001}) - b_{A,Ea}} = Z_{0.999}(\mu_{T_0}, \sigma_{T_0})$, where $Z_{0.999}(\mu_{T_0}, \sigma_{T_0})$ is a quantile of $N(\mu_{T_0}, \sigma_{T_0})$.

This leads to $tc_{0.001} = 10^{\frac{a_{A,Ea}}{Z_{0.999}(\mu_{T_0}, \sigma_{T_0})} + b_{A,Ea}}$. Likewise, $tc_{0.999} = 10^{\frac{a_{A,Ea}}{Z_{0.001}(\mu_{T_0}, \sigma_{T_0})} + b_{A,Ea}}$.

The values of the probability density *f(tc)* will then be computed for *n* = 1000 values of *tc* in the interval [$tc_{0.001}$; $tc_{0.999}$].
Given the joint distribution of the parameters (*A*, *Ea*) $f_{A,Ea}$, the distribution of *tc* is given by $f(tc) = \int_{A\in[A_{min},A_{max}],Ea\in[Ea_{min},Ea_{max}]} f(tc|A,Ea) f_{A,Ea} dA dEa$.
whereby $f_{A,Ea}$ might be either a prior distribution or a posterior distribution after taking the species concentration profiles into account.
We compute it on the interval $[min_{A,Ea}(tc(A,Ea)), max_{A,Ea}(tc(A,Ea))]$.

The true probability distribution of *tc* for A_0 and Ea_0 is as follows:

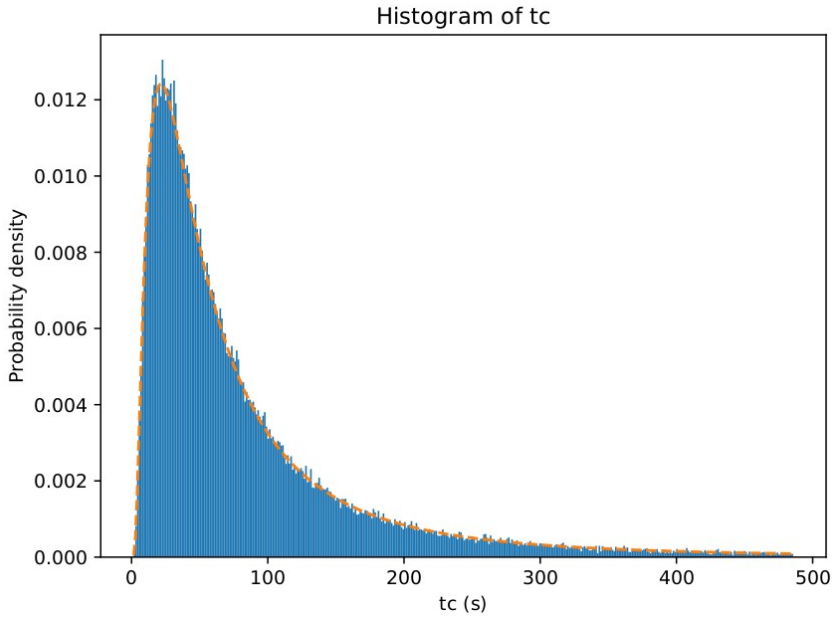
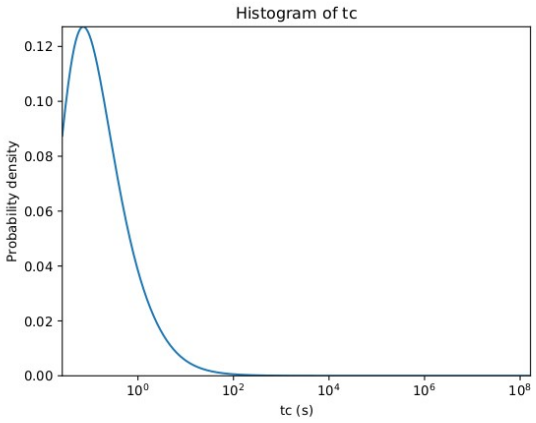


Figure 1: Probability distribution for A_0 = **8.6E+11** *and Ea_0 =* **30**

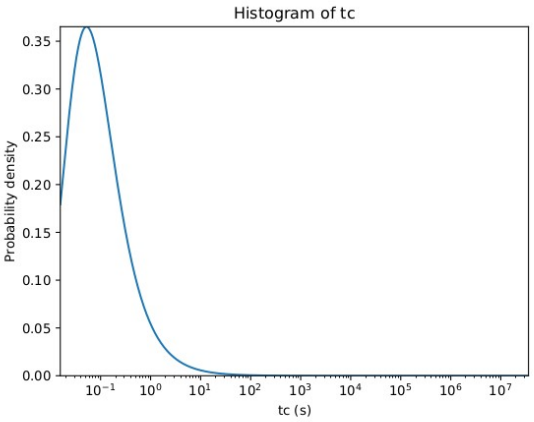
8.6E+11 - 30	tc_{25}	tc_{50}	tc_{75}	$p_{critical}$
Analytical	28.86	55.08	108.79	0.2634
Numerical	28.94	54.94	107.72	0.2625

In what follows, *f(tc)* and its main features will be given in Situation A, B, C and D.

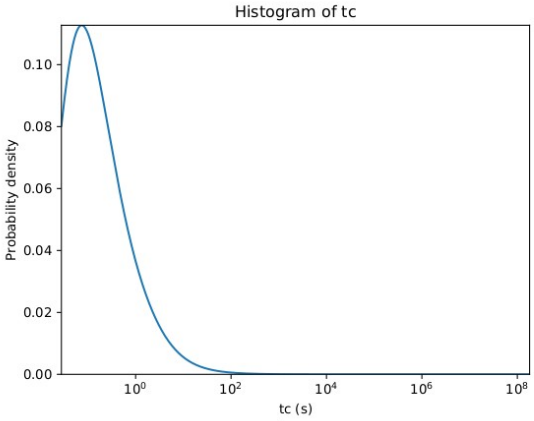
Situation A



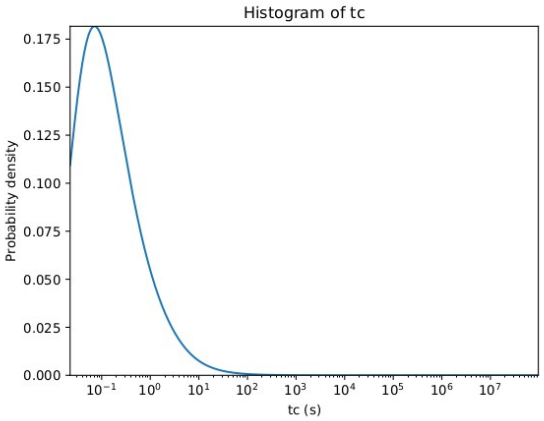
$\mathbf{f_{A,1}}$



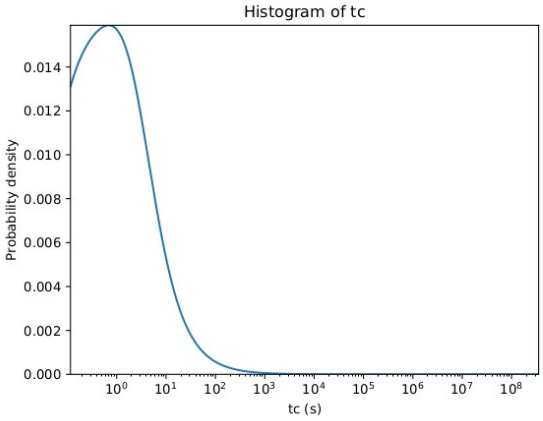
$\mathbf{f_{A,2}}$



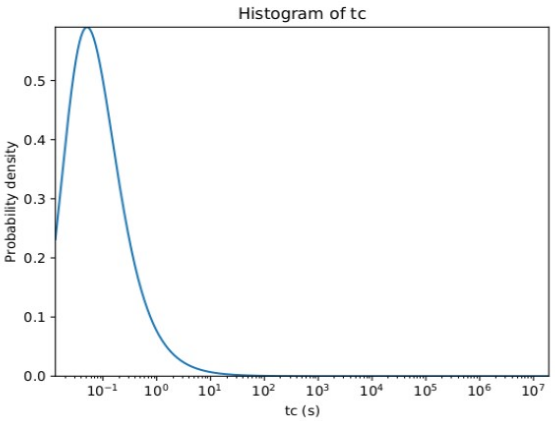
$\mathbf{f_{A,3}}$



$\mathbf{f_{A,4}}$



$\mathbf{f_{A,5}}$

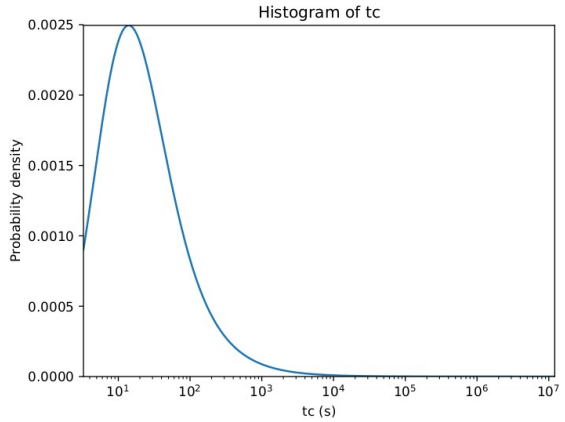


$\mathbf{f_{A,6}}$

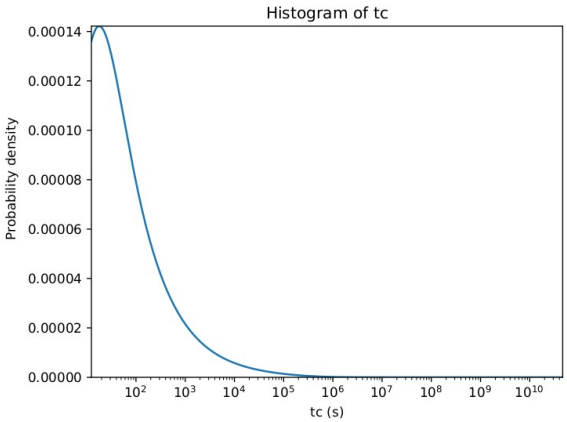
Prior	tc_{25} (s)	tc_{50} (s)	tc_{75} (s)	$p_{critical}$
A_0, Ea_0	**28.86**	**55.08**	**108.79**	**0.2634**
$f_{0,1}$	34.29	2801.66	2.34E+05	0.2424
$f_{0,2}$	7.85	649.06	5.45E+04	0.3260
$f_{0,3}$	39.22	3201.74	2.68E+05	0.2348
$f_{0,4}$	9.39	356.02	4.00E+04	0.3360
$f_{0,5}$	149.45	1.21E+04	9.99E+05	0.1585
$f_{0,6}$	1.80	69.51	8019.06	0.4476
[min; max]	**[1.80; 149.45]**	**[69.51;1.21E+04]**	**[8019.06;9.99E+05]**	**[0.1585;0.4476]**

As could be expected, *f(tc)* and its features widely differ if the non-updated priors $f_{0,i}(A,Ea)$ are used and none of the *f(tc)* are close to the true distribution obtained with A_0 and Ea_0.

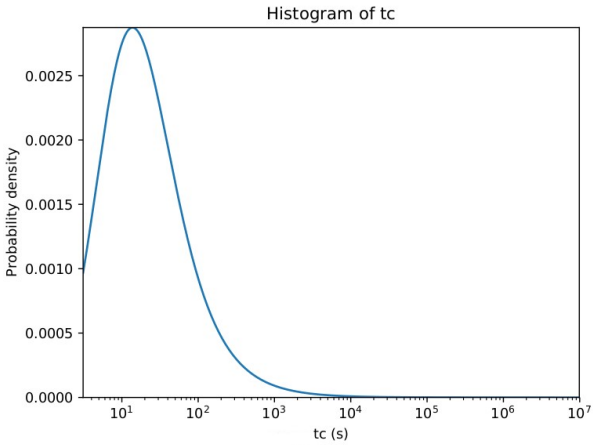
Situation B



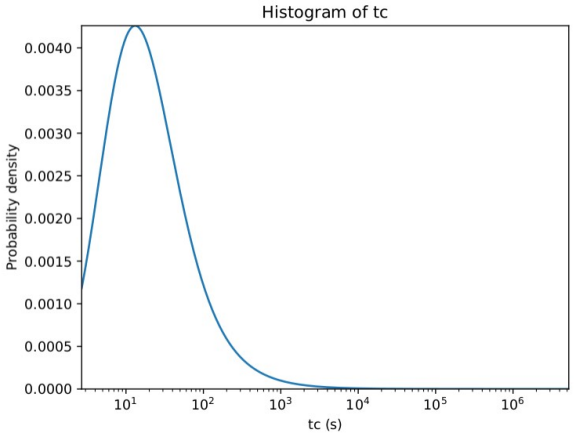
$\mathbf{f_{B,1}}$



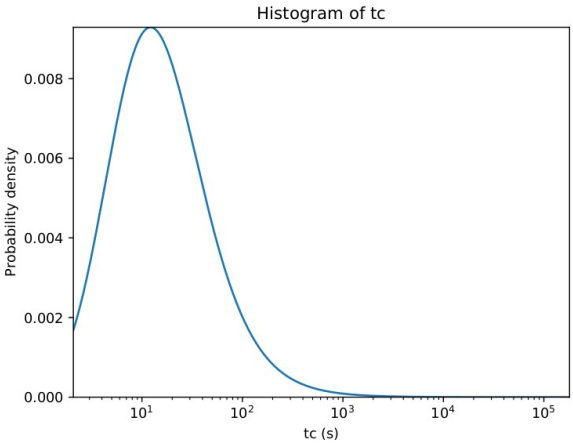
$\mathbf{f_{B,2}}$



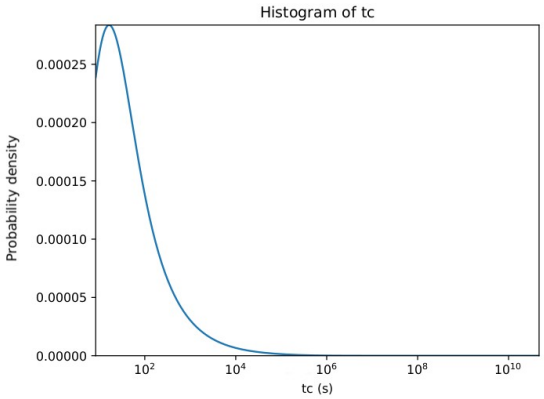
$\mathbf{f_{B,3}}$



$\mathbf{f_{B,4}}$



$\mathbf{f_{B,5}}$

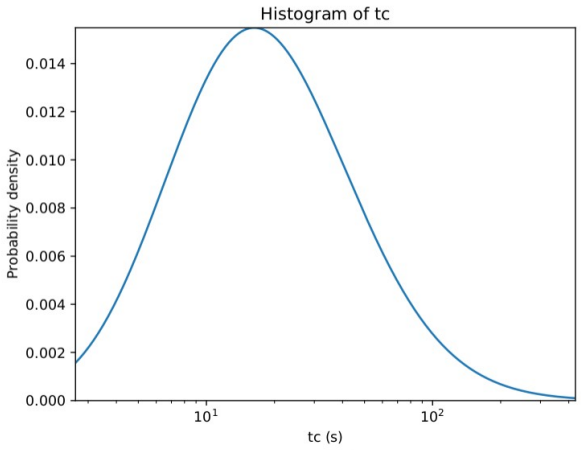


$\mathbf{f_{B,6}}$

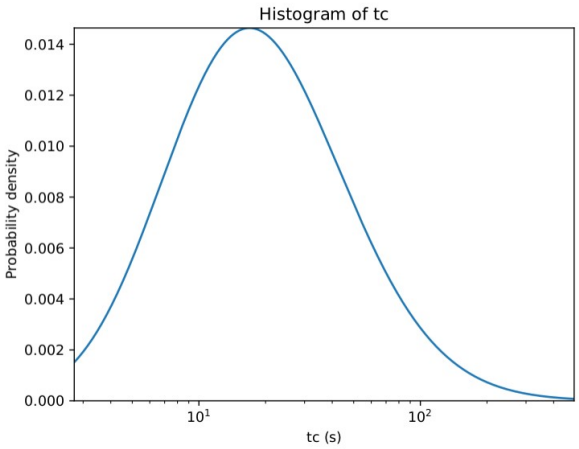
Posterior	tc_{25}	tc_{50}	tc_{75}	$p_{critical}$
A_0, Ea_0	**28.86**	**55.08**	**108.79**	**0.2634**
$f_{B,1}$	322.64	5667.50	101985.48	0.0600
$f_{B,2}$	46135.45	267339.01	1194153.91	0.0034
$f_{B,3}$	234.54	3630.43	71792.60	0.0690
$f_{B,4}$	108.43	1035.65	21514.09	0.1018
$f_{B,5}$	35.05	114.69	604.56	0.2183
$f_{B,6}$	26002.85	192780.68	959438.88	0.0069
[min;max]	**[35.05;4.61E+04]**	**[114.69;2.67E+05]**	**[604.56;1.19E+06]**	**[3.4E-03;0.2183]**

As can be seen in **Appendix C: virtual experimental data** (fischer21c-supp.pdf), the use of only one experiment at a single temperature is not enough to determine the two kinetic parameters *A* and *Ea* at the same time. A very large number of (*A*, *Ea*) values are equally probable. Consequently, the probability distribution of *tc* obtained with most posteriors tends to be far off from the real one and the computed distributions themselves strongly disagree with one another.

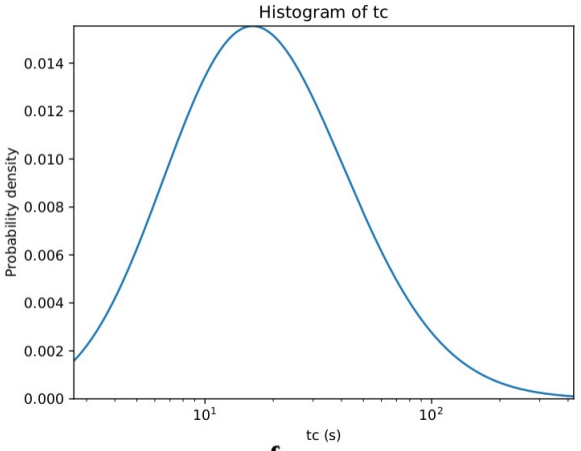
Situation C



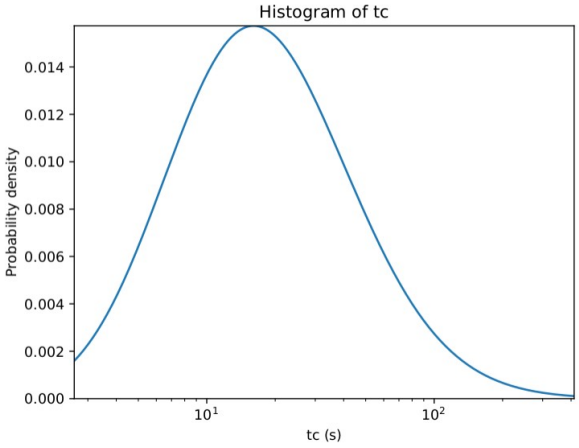
$\mathbf{f_{C,1}}$



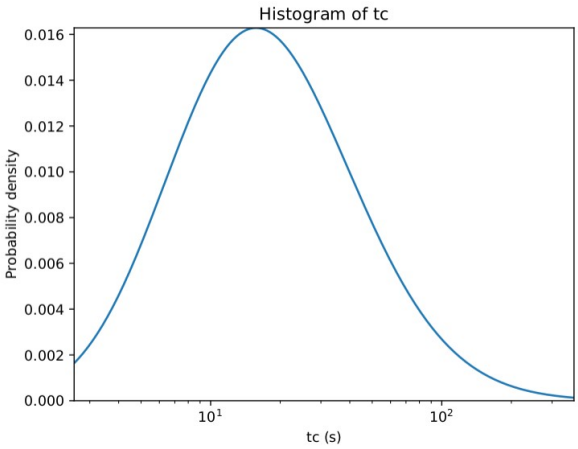
$\mathbf{f_{C,2}}$



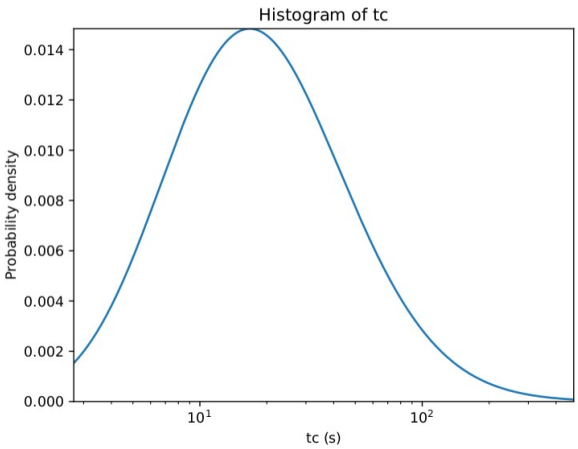
$\mathbf{f_{C,3}}$



$\mathbf{f_{C,4}}$



$\mathbf{f_{C,5}}$

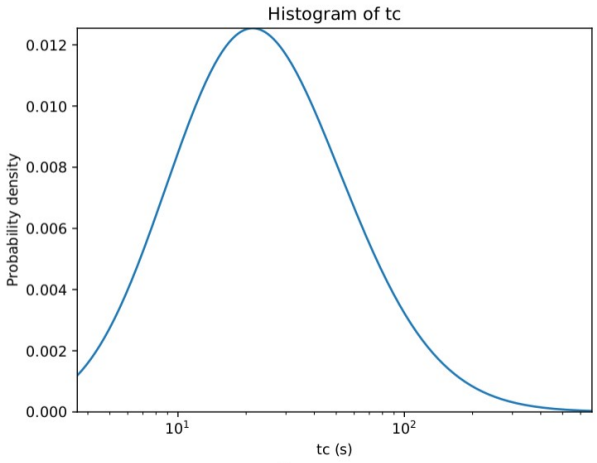


$\mathbf{f_{C,6}}$

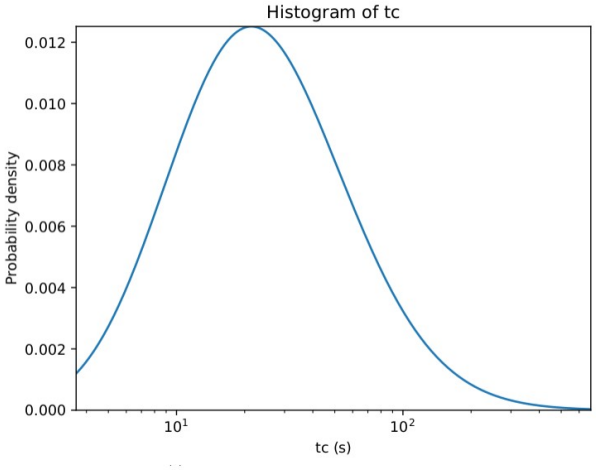
Posterior	tc_{25}	tc_{50}	tc_{75}	$p_{critical}$
A_0, Ea_0	**28.86**	**55.08**	**108.79**	**0.2634**
$f_{C,1}$	22.56	44.03	88.39	0.3508
$f_{C,2}$	23.74	46.62	94.41	0.3311
$f_{C,3}$	22.49	43.86	87.99	0.3521
$f_{C,4}$	22.24	43.32	86.77	0.3566
$f_{C,5}$	21.57	41.83	83.22	0.3689
$f_{C,6}$	23.45	46.00	93.02	0.3358
[min;max]	**[21.57;23.74]**	**[41.83;46.62]**	**[83.22;94.41]**	**[0.3311;0.3689]**

Situation C corresponds to two noisy experiments at two different temperatures.
As shown in **Appendix C: virtual experimental data** (fischer21c-supp.pdf), the region of probable values of (A, Ea) is considerably smaller than in situation B. While the differences between the delay times and $p_{critical}$ are considerably smaller than in situation A and B, the delay times are systematically under-predicted whereas $p_{critical}$ is systematically over-predicted. This could be an artefact of the choice of the six prior distributions and could possibly be avoided by choosing a parametric family of priors instead.

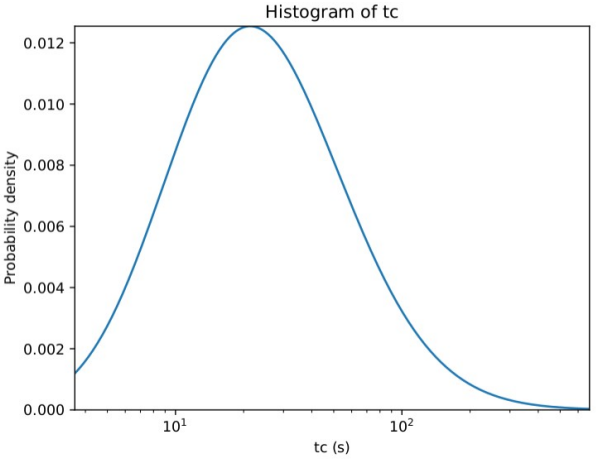
Situation D



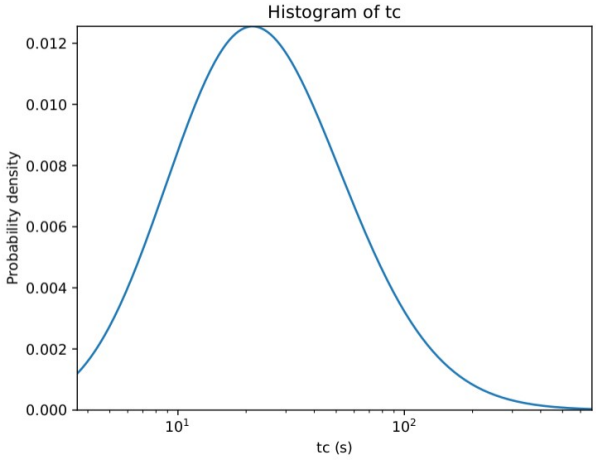
$\mathbf{f_{D,1}}$



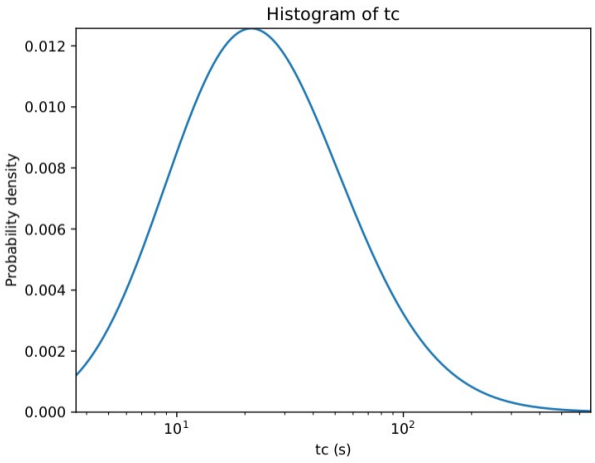
$\mathbf{f_{D,2}}$



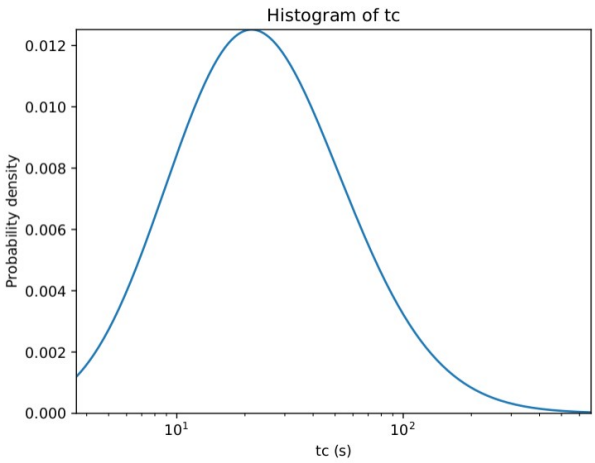
$\mathbf{f_{D,3}}$



$\mathbf{f_{D,4}}$



$\mathbf{f_{D,5}}$



$\mathbf{f_{D,6}}$

Posterior	tc_{25}	tc_{50}	tc_{75}	$p_{critical}$
A_0, Ea_0	**28.86**	**55.08**	**108.79**	**0.2634**
$f_{D,1}$	28.55	54.53	107.77	0.2672
$f_{D,2}$	28.61	54.65	108.03	0.2664
$f_{D,3}$	28.54	54.52	107.75	0.2672
$f_{D,4}$	28.53	54.49	107.70	0.2674
$f_{D,5}$	28.48	54.40	107.51	0.2680
$f_{D,6}$	28.60	54.62	107.98	0.2666
[min;max]	**[28.48;28.61]**	**[54.40;54.65]**	**[107.51;108.03]**	**[0.2664;0.2680]**

In that situation, the six posteriors were computed from four experiments at four different temperatures with a low relative standard deviation σ_r = 6%. The quantiles and $p_{critical}$ are almost the same. The very slight under-prediction of the delay times and over-prediction of $p_{critical}$ might stem from the numerical approximations used to compute t_c or from the fact that no parametric family of prior distributions was considered for this study.