“Petroleum generation kinetics: Single versus multiple heating-ramp open-system pyrolysis”


Discussion

Douglas W. Waples

Sirius Exploration Geochemistry, 9299 William Cody Drive, Evergreen, CO 80439

d.waples@siriusxgc.com.

Peters et al. (2015) have criticized the one-run kinetics method that my colleagues and I have developed and promoted (Waples et al., 2002, 2010; Waples and Nowaczewski, 2013¹), and claim further that source-rock kinetics derived by the standard multirun methods are superior. Specifically, they state that A factors determined mathematically from multirun kinetics are significantly more accurate than those obtained using other methods, and that those A factors and the associated activation-energy (Ea) distributions are sufficiently accurate in an absolute sense to be used, with confidence and without any further quality control, in modeling hydrocarbon generation. I disagree strongly with those various statements.

¹ Contrary to the citation, this paper was never published. Because of a communication error, it was cited by Peters et al. (2015) in this format. That manuscript has been updated and replaced by Waples (submitted) after Vincent Nowaczewski requested that his name be withdrawn for personal reasons. The unpublished version cited by Peters et al. is available at http://siriusxgc.com/kinetics-papers/. For non-specialists interested in the subject, I recommend instead the later Waples (submitted) version, which will also be available at that same website until it is published.
Any discussion of the kinetics of hydrocarbon generation must begin with an understanding that the parallel first-order kinetic description used by almost everyone is simply a convenient construct, and is clearly not a mechanistic description of what happens in nature (e.g., Stainforth, 2009). Most researchers in this field believe that the reactions involved in hydrocarbon generation are mainly chain reactions rather than simple decompositions; that reaction schemes do not involve a group of discrete parallel processes; and that generation yields a variety of intermediate products on the way to final products (e.g., Stainforth, 2009; Tang and Ma, 2009). We therefore seek a useful kinetic system, rather than a correct one, and we must be prepared to make compromises and simplifications to achieve this goal. That said, not all compromises and simplifications are equally valid or acceptable.

The parallel first-order kinetic model has been remarkably successful for three decades in modeling the evolution of the S2 peak during Rock-Eval-type pyrolysis of the great majority of kerogens, although exceptions have been noted where the S2 peak is narrower than first-order kinetics can explain (e.g., Burnham et al., 1996; Stainforth, 2009). Because generation of the material comprising the S2 peak is very widely considered to be an excellent proxy for hydrocarbon generation, the parallel first-order kinetic model, when properly adapted to fit both laboratory conditions and geologic conditions, provides a very useful way to make good predictions about the status and timing of hydrocarbon generation.

The most-important differences between my own perspective and that of Peters and his collaborators lie in how to define an appropriate value for a single Arrhenius (A) factor for the
entire set of parallel reactions that are derived by mathematical workup of pyrolysis data. Their method finds the combination of $A$ and a set of activation energies ($E_a$) that gives the absolutely best mathematical fit between pyrolysis curves measured in the laboratory at several (usually three to five) different heating rates, and synthetic pyrolysis curves calculated at those same heating rates. No *a priori* constraints are placed on the values of $A$ or $E_a$; those constants are determined mathematically without invoking any scientific knowledge. Multiple runs at different heating rates are necessary in order to separate the linear dependence of the generation rate on the $A$ factor from the exponential dependence on the $E_a$ distribution.

Workup of the raw pyrolysis data is different using the one-run method. The value of $A$ must be pre-specified, since the influences of $A$ and $E_a$ cannot be separated when using only a single pyrolysis run. Once $A$ is specified, however, a full $E_a$ distribution can be determined from a single lab analysis. Waples et al. (2010), Waples (submitted), and Waples and Nowaczewski (2013) have discussed the benefits and assumptions of the one-run method.

The fundamental problem that has surfaced during the debate about one-run kinetics versus multirun kinetics is thus how one can best obtain a reliable value for $A$. Once $A$ has been determined, it is a simple matter to find an $E_a$ distribution that fits the laboratory data. There are three approaches to determining the value of $A$ for any particular kerogen. Most commonly, multirun kinetics provides a value for $A$ automatically. With one-run kinetics, in contrast, the user must supply a value for $A$, and thus he or she must have an independent source of
information that enables this choice. The third possibility is to use a theoretical approach based on chemical kinetics and the thermodynamics of transition states.

Unfortunately, a purely theoretical approach cannot achieve this goal, because although A and Ea are understandable and predictable for simple chemical reactions (Benson, 1976), any application of that information to chain reactions requires a detailed knowledge, including chain lengths, of all the reactions in the system. Such knowledge is clearly impossible for complex and heterogeneous systems in nature, such as hydrocarbon generation.

The standard multirun method of determining A has been severely criticized, beginning 25 years ago, by various researchers (Lakshmanan et al., 1991; Nielsen and Dahl, 1991; Stainforth, 2009; Nordeng, 2015). To my knowledge, these criticisms have never been refuted, and must be taken very seriously. The basic technical problem with multirun kinetics is that although the fundamental mathematical idea behind the method is valid, the range of heating rates available from laboratory analyses is not large enough to overcome measurement errors that occur during the pyrolysis runs, and thus cannot provide an accurate value for A, or a corresponding Ea distribution. Rather than finding a combination of A and Ea that gives a substantially better fit to the pyrolysis data than any other combination, we always find that a wide range of A-Ea combinations will give fits to laboratory data that are statistically almost indistinguishable. For example, as Figure 1 shows, for most samples, over a range of logA values from about 10 to 17 (that is, seven orders of magnitude variation in A), the residual error in the quality of fit to the pyrolysis data varies by less than a factor of two. Figure 2 shows that the fit to the data for one
of those samples using the freely chosen A factor of $10^{13} \text{ s}^{-1}$ is not visually much different from that obtained using fixed A factors of $10^{11}$ and $10^{17} \text{ s}^{-1}$. Results of this type, which are extremely common, cannot possibly yield confident values for the A factor.

The most vociferous, comprehensive, and telling criticism of the multirun kinetic method has come from Stainforth (2009), who discusses, among other topics, the so-called compensation effect. In the field of source-rock kinetics, this term refers to the fact that the observed rate of a chemical reaction can be explained with good accuracy by many combinations of A and Ea, in which an increase in A (which accelerates the reaction) requires an increase in Ea, which slows the reaction. Stainforth distinguishes between “real” compensation effects (which correspond to true natural variation in A-Ea combinations) and “false” compensation effects, which are artifacts arising from the inability of the multirun data-workup methods to choose the true A factor.

Compensation effects in petroleum generation kinetics have been discussed numerous times before … but just how false they typically are, and how detrimental to the extrapolation to geological conditions, may not have been fully appreciated.

False compensation effects are so widespread in the literature of petroleum generation that almost any published data set … can be used to illustrate their characteristics.
He also points out the unrealistic thermodynamic implications of applying A-Ea combinations derived from purely mathematical fits to pyrolysis data:

"... the widely ranging A factors imply activation entropies ranging from -110 to +90 J/K mol, which are atypically large in magnitude for the overall decomposition of organic materials ... it is strange that the overall activation entropies for peak oil generation should range from negative to positive rather than being consistently one or the other."

Another related indication of problems with the automated multirun method is when the overall reaction-rate constants from laboratory analyses of two samples are very similar, but the A and Ea values determined from those constants are very different (Stainforth, 2009). In fact, when the data-workup software chooses the very-best mathematical fit from among these myriad very-similar solutions, its choice is very likely to be influenced mainly by factors related to analytical uncertainty, rather than to real variation in the characteristics of the kerogen. Ironically, Stainforth’s point is illustrated very well but unintentionally by the data of Peters et al. (2015) (their Table 3), as well as by data presented by Nordeng (2015). Using the data of Peters et al. (2015), and limiting our discussion conservatively to those examples that were calculated using at least three different heating rates, with at least one rate lower than or equal to 5°C/minute and at least one rate greater than or equal to 30°C/minute, the freely selected A factors for the Kimmeridge Clay sample vary by a factor of 7.7, from $3.16 \times 10^{13}$ s$^{-1}$ to $2.42 \times 10^{14}$ s$^{-1}$. Using the same criteria for the Monterey Shale sample, the freely selected A factors vary by a
factor of more than 12, from $6.39 \times 10^{12}$ s$^{-1}$ to $7.78 \times 10^{13}$ s$^{-1}$. Using the 1-2-3 rule cited by Peters et al. (2015), these uncertainties in A will lead to uncertainties of about 3 kcal in Ea and about 9°C in generation temperature under geologic conditions, and slightly greater uncertainties for the Monterey. Finally, these variations are even greater if one applies less-stringent criteria to the range or number of heating rates.

In summary, the data of Peters et al. (2015) contradict their conclusion that multirun kinetics give highly confident or unique A factors and corresponding Ea distributions. On the contrary, those data confirm the beliefs of myself and several other workers that the multirun method is fraught with potentially dangerous uncertainties that cannot be easily overcome, and should definitely not be ignored.

Stainforth comments further on the consequences of using multirun kinetics where A and Ea combinations are chosen automatically:

Such false compensation effects between E and A parameters result in very poor extrapolations of measured laboratory rate data, both to higher laboratory temperatures and, more seriously, to the much lower temperatures of petroleum generation under geological conditions...
It is worth noting that one of the data sets Stainforth used to illustrate false compensation effects comes from Peters et al. (2006), which is one of the main data sets Peters et al. (2015) used to support their arguments about appropriate values for A factors.

Alan Burnham (personal communication, 2014) has suggested that much of the variation in A factors reported in early studies was the result of poor temperature control in early kinetics determinations, and agrees with me that the natural variation in A is less than early work suggested. To test the hypothesis that changes in technology have resulted in changes in measured A factors, I have compared in Figure 3 the published A factors obtained between about 1988 and 2013, with those obtained only since the year 2000. While there are fewer recent A values near the extreme upper and lower limits, the most-extreme high and low values were obtained recently. The standard deviation in both populations is virtually identical, and the average A factor for the recent data set is slightly more than two times that for the complete data set. I conclude that these data do not clearly support Burnham’s suggestion that improved technology has resulted in better kinetic parameters and a narrower range of A factors. Consequently, these data therefore suggest that improvements in technology since 1988 have not diminished the role of the compensation effect in controlling A factors obtained using the standard multirun method. It is still possible, however, that some early data could have been affected by poor technology.

If we admit that theory is incapable of providing specific values for A, and if we acknowledge that the multirun method as necessarily employed does not consistently provide acceptable A
factors, we have eliminated all possible methods for determining A except empirical ones. I suggest here three empirical methods for estimating values for A. At this stage of our knowledge we must recognize that this approach is imperfect, and that our final choice of A will therefore always have an associated uncertainty.

The first method is that suggested by Waples (submitted) and Waples and Nowaczewski (2013). A single value for A would be selected from the distribution of A factors determined by allowing A and Ea to vary freely (Figure 3, top). Median and mean values for logA both yield a value for A near $1.6 \times 10^{14}$ s$^{-1}$. This value rounds up to the value of $2 \times 10^{14}$ s$^{-1}$ that I normally use. The rationale for using this method is that the true value for A for any kerogen should fall in the middle of the symmetrical compensation ellipse that various workers have described (Lakshmanan et al., 1991; Nielsen and Dahl, 1991; Nordeng, 2015).

The second method is to choose a value for A that yields a good correspondence between calculated hydrocarbon generation and calculated vitrinite reflectance. All petroleum geochemists and basin modelers have in their minds a definition of the hydrocarbon-generation window or oil-generation window in terms of Ro, although the specific opinions about the correlation between Ro and generation will vary from worker to worker, as well as with kerogen type. In spite of these real or imagined variations, however, we can put some highly confident limits on the acceptable relationships between Ro and level of generation.
Most geochemists and modelers would agree, I believe, with the statement that the true Ro value at peak hydrocarbon generation (that is, Transformation Ratio [TR] = 0.5) for all kerogens (except possibly very unusual ones, such as resinites or perhaps some high-sulfur kerogens) would fall somewhere between 0.75% Ro and 1.2% Ro. Figure 4 shows calculated Ro as a function of calculated TR for three typical published kerogens (representative of Types I, II, and III) using a variety of A-Ea combinations for each kerogen. The A factors used for each kerogen range from $10^{12}$ s$^{-1}$ to $10^{17}$ s$^{-1}$, and include the published A factor and Ea distribution. The unpublished A-Ea combinations were calculated using a simple algebraic adjustment that forces the adjusted kinetics to give the same overall reaction rate as the original kinetics at laboratory temperatures. The geologic heating rate used for each calculation was 3°C per million years.

For Type I kerogen, the A factor in Figure 4 required to yield a TR of 0.5 at 0.75% Ro is $2.1*10^{13}$ s$^{-1}$, while that required to yield the same TR at 1.2% Ro is $3*10^{15}$ s$^{-1}$. These two values for A therefore bracket the maximum acceptable range for Type I kerogen. For Type II kerogen the range is from $1.3*10^{13}$ s$^{-1}$ to $3.2*10^{15}$ s$^{-1}$, while for Type III kerogen it is $1.9*10^{13}$ s$^{-1}$ to $2*10^{16}$ s$^{-1}$.

For Type I kerogen the maximum A is 140 times that of the minimum A, while for Type II kerogen the ratio is 250, and for Type III kerogen it is about 1000. The possible variations in A represented by all these ratios are significantly smaller than the four orders of magnitude suggested by Peters et al. (2015).

However, if we use instead my personal preference for a minimum Ro of 0.9% and a maximum Ro of 1.1% at peak hydrocarbon generation in most kerogens, the range of variation in A factors
obtained from Figure 4 becomes much smaller. The range for Type I kerogen becomes $6 \times 10^{13} \text{ s}^{-1}$ to $4 \times 10^{14} \text{ s}^{-1}$ (ratio = 7); for Type II kerogen it becomes $3.5 \times 10^{14} \text{ s}^{-1}$ to $8 \times 10^{14} \text{ s}^{-1}$ (ratio = 2.5); and for Type III kerogen it becomes $3.1 \times 10^{14} \text{ s}^{-1}$ to $2.2 \times 10^{15} \text{ s}^{-1}$ (ratio = 7). All the ratios derived for the narrow window for peak generation are close to the very-narrow range ($8 \times 10^{13} \text{ s}^{-1}$ to $4 \times 10^{14} \text{ s}^{-1}$; ratio = 5) suggested by Waples and Nowaczewski (2013) for kerogens in general. However, in light of the results presented above, and to be conservatively all-inclusive, it might be reasonable to increase the suggested maximum range of allowable A values for all types of kerogen to be from about $6 \times 10^{13} \text{ s}^{-1}$ to $2 \times 10^{15} \text{ s}^{-1}$ (ratio = 33).

It is also worth noting that the value of $1.6 \times 10^{14} \text{ s}^{-1}$ obtained from the mean and median of the total population in Figure 3 falls within or very near all these ranges. Therefore, our two empirical methods yield results that are consistent with each other, but the second method gives us an estimate of the range of plausible A values as well as just the best estimate.

A third empirical method for specifying A and the variability in A is also available, but it is not illustrated in this discussion because it is very similar to the method using calculated Ro values. In this third method we would calculate the Transformation Ratio for a sample from its measured Hydrogen Index, as described by Waples and Tobey (2015), and find the range of A values that is consistent with both our best estimate of the TR, and our uncertainty in our calculated TR. Of course, Ro and TR could be used together to give even better resolution than either parameter alone.
In their paper Peters et al. (2015) repeatedly referred to the “error” associated with using kinetics derived using the one-run method. Since I have shown above that both multirun and one-run kinetics have uncertainties, I believe it is fairer to both sides, and also more accurate, to call the discrepancy between results obtained using the two types of kinetics a “difference” rather than an error.

Finally, there are two additional minor errors in the manuscript that misrepresent my views. Waples and Nowaczewski (2013) did indeed cite Rock-Eval as an appropriate tool for deriving one-run kinetics, but never stated that the reported Rock-Eval temperatures should be used without correction, as Peters et al. (2015) imply.

More importantly, Peters et al. (2015) stated several times that I had indicated that a “universal” value for A should be used for all kerogens. That this statement is untrue is clearly contradicted by the statement of Waples and Nowaczewski (2013) that

\[ \ldots \text{we believe that the great majority of A factors for hydrocarbon generation from different kerogens can range from about } 8 \times 10^{13} \text{ s}^{-1} \text{ to } 4 \times 10^{14} \text{ s}^{-1}, \text{ with a dominance near } 2 \times 10^{14} \text{ s}^{-1}. \]

At the present state of our knowledge I do in fact prefer to use the same value of A for all kerogens, mainly for the dual conveniences of not having to guess A factors for specific kerogens, and of being able to see at a glance which kerogen will generate earlier. However, I
have no objection to anyone who would prefer to use different A factors for different kerogens, provided those A factors yield defensible results when applied at geological temperatures.

In summary, the conclusion of Peters et al. (2015) that the actual range of A factors for hydrocarbon generation can span a range of four orders of magnitude is not supported by existing information. First, the chemical reactions they cited in support of this idea may have little in common with hydrocarbon generation taking place in high-pressure aqueous subsurface environments. Second, the discussion above shows clearly that a large range of A factors will lead to conflicts about the width of the hydrocarbon-generation window as defined using other geochemical parameters. Finally, their statement that the variation of four orders of magnitude in A factors that they observed in data derived from their multirun kinetics proves that A factors can vary by that amount is a classic example of circular reasoning. Their measured values cannot be used to prove that the method that generated those values is valid.

Finally, in spite of the many difficulties he noted in determining reliable kinetic parameters, Stainforth is guardedly optimistic about the value of source-rock kinetics, provided that one recognizes the danger of false compensation effects and finds a way to eliminate those effects:

Extrapolation of the kinetics from laboratory to geological conditions is demonstrably successful, so long as false compensation is avoided or factored out.

References


Nordeng, S. H., 2015, Compensating for the Compensation Effect Using Simulated and Experimental Kinetics From the Bakken and Red River Formations, Williston Basin, North Dakota:


Tang, Y. and Q. Ma, 2009, Integrating Advanced Geochemistry into Basin Modeling: Applications for Exploration and Production Risk Analysis:


Figure 1. Relative residual error (relative to a value of 1.0 for the lowest error for each sample) for seven samples, each run at five different heating rates (2, 5, 15, 30, and 50°C/min). Fits for each sample were achieved by fixing log A (s⁻¹) to integral values from 10 through 20 (that is, A from 10¹⁰ s⁻¹ to 10²⁰ s⁻¹). Also shown are the data from the best fit for each sample, determined by allowing A and Ea to vary freely during standard multirun kinetic analysis. Sample identities are TA = Rudeis, Miocene of the Gulf of Suez, Egypt; NO = Thebes, Eocene of the Gulf of Suez, Egypt; LE = Chekka, Paleocene of Lebanon; ES = Umm Ghudran?, Maastrichtian, Jordan; OS = Duwi, Campanian, Eastern Desert of Egypt; GE = Brown Limestone, Coniacian, Gulf of Suez of Egypt; RA = Abu Roash “F”, Turonian, Western Desert of Egypt. Data courtesy of StratoChem Services.

Figure 2. Fit between measured pyrolysis curves obtained at heating rates of 2, 5, 15, 30, and 50°C/min for sample RA (Abu Roash “F”, Turonian of the Western Desert of Egypt). Top: Fixed A = 10¹¹ s⁻¹. Middle: A factor of 1.9973*10¹³ s⁻¹ chosen by Kinetics05 software to provide the best mathematical fit to the data. Bottom: Fixed A = 10¹⁷ s⁻¹. Data courtesy of StratoChem Services.

Figure 3. Histograms of the log base 10 of A factors for the kinetics of total hydrocarbon generation, determined using the standard multirun method in which A and the Ea distribution are allowed to vary freely. Top: All published A factors from about 1988 through 2013. Bottom: Data for samples in the upper histogram that were published from 2000 to 2013.
Figure 4. Calculated Ro vs. calculated TR for typical Type I (top), II (middle), and III (bottom) kerogens using the published A-Ea combination (black line) and several adjusted A-Ea combinations for each. Type I from Peters et al. (2006). Type II from Schenck and Dieckmann (2004). Type III from Di Primio and Horsfield (2006). Vertical black line represents TR= 0.5 (peak hydrocarbon generation). Gray shaded area represents range of plausible Ro values when TR = 0.5. Legend labels represent the log of the A factor expressed in s^{-1}. 
Figure 1.
Figure 2.
Figure 3.
Figure 4.