# The importance of process in modern tectonostratigraphy and regional geology

**Peter Lunt[1], Xiwu Luan[2]**

[1]Stratos Energy Advisors, Kuala Lumpur, Malaysia.
[2]Shandong University of Science and Technology, Qingdao, China

## ABSTRACT

In its simplest form geology is a history, which appears to only require the compilation of a chronicle. However, our data is fragmentary, and aspects such as the dimension of time and the depositional setting of sediments are obscured, especially the correlation of time over wide areas and determining rates of change. As a result, geology cannot advance just by accumulating observations, expecting that a credible narrative will inevitably emerge. There must be reality checks on the proposed history to avoid confirmation bias towards over-simple expectations. In a region suspected to be complicated and, in many ways, unique we require an abductive investigative technique to reconstruct its history. This means to force testing through cross-checking independent but related data types. Such inter-disciplinary testing can produce a rigorous framework, even to reconstruct special situations not accommodated by ideas models. This essay examines three topics necessary for such evidence-based investigation. The first is to investigate and document the reliability of observations (like the "error-bars" required in hard-sciences). Secondly is the shift from model-based to evidence-based processes (from deductive to abductive reasoning). Thirdly is the need to consider if conclusions are "significant" - i.e., is there confidence that an interpretation would be repeatable by independent workers, as well as being distinct from background variability in data. It is proposed that we must acknowledge the replication crisis highlighted in the past two decades in other sciences by considering how we work in the complex geology of SE Asia, to prevent a similar validation crisis undermining the value of the science here.

***Keywords:*** *tectonostratigraphy, regional geology, abductive investigative technique*



*Manuscript received: 6 Oct 2023, revised manuscript received: 5 Dec 2023, final acceptance: 18 Dec 2023. DOI: 10.51835/bsed.2024.49.3.430*

## INTRODUCTION

This is the second in a series of related essays on the general properties of geological studies in the dynamic setting of SE Asia. The first examined the application of names in a science rich in nouns and jargon (Lunt and Luan, 2023).

There seems to be an unstated assumption that as more data is gathered it will inevitably trend towards better conclusions. However, other sciences have raised an alarm on this topic, so here we attempt to outline concerns on the possible misuse of process in Asian geology. The work of Ioannidis (2005) led to recognition of the so-called "replication crisis" in experimental sciences, as documented by Baker (2016) and discussed in general magazines (e.g., New Scientist, 6 April 2022). The term p-hacking came out of this reflection on mostly accidental misuse of processes, as did the related term HARKing (Hypothesizing After Results are Known; Kerr, 1998), which highlights a combination of confirmation bias and cherry-picking to match a new model or hypothesis on the assumption that the new hypothesis (or technology) must be right. A form of HARKing is sometimes relevant in exploration geology when new technology presents visually impressive, often expensive results, and data is then arranged to support a positive interpretation. Instead, it might have been better asking if the whole process might have been a waste of resources to reach nothing outside a null-hypothesis, or beyond a simple tautology.

Previous examples of the examination of process in stratigraphic geology include the discussion of the hermeneutic cycle by Miall and Miall (2004), which were attempts to balance the empirical (inductive) stages of reasoning with deductive stages from a theoretical model or a current paradigm/framework. As both approaches can be flawed, the hermeneutic cycle is a process of repeatedly moving between the two contrasting ways of reasoning to bring about meaningful, self-checking application and advance. What is proposed here is an adaptation of this approach for tectonically active areas, which exploits the excellent biostratigraphy in SE Asia from tropical Cenozoic micropalaeontology. As noted by Miall and Miall "*But if there is any lesson that two centuries of geological investigation since the days of William Smith have taught us, it is that* ***biostratigraphy is the ultimate arbiter of chronostratigraphic correlations****. The literature is full of lithostratigraphic correlation schemes that have failed because of insufficient attention to biostratigraphy. … Blindly correlating stratigraphic events to the outdated onlap-offlap curve of Haq et al. (1987, 1988) without determining whether biostratigraphic data support their correlations, continues the trend of poor science.*" (their emphasis). In SE Asia the excellent micropalaeontology means we can also add tests from examining facies shifts and Walther's Law, an application that greatly extends the usual tools of biostratigraphic correlation and age dating.

It is proposed that SE Asia requires a repeatedly inductive hermeneutic cycle. That is, the excellent facies and age data is used along with lithostratigraphy, seismic and other data, to induce new concepts, but this is followed by a stage of extending induction into related disciplines. This is a deliberate search for tests of concepts in the adjacent data sets, also known as abduction; or the forcing of resolution through tests (abducere means to lead or force through). Abductive reasoning is strictly a ranking of plausibility between

what geologists would call multiple working hypotheses (the latter being an idea developed by the 19th century geologist T. C. Chamberlin). Once a single hypothesis stands out as the most reasonable, it becomes the updated geological framework (a multi-disciplinary hypothesis; distinct from a conceptual model) and this framework can be temporarily used for deduction. However, the cycle of investigation is then repeated by deliberately obtaining new data that is designed to test, and not just confirm, the framework. This process of deliberately forcing integration of related disciplines will both test existing concepts as well as potentially highlight any anomalies that can lead to paradigm shifts.

An example from the region of an entrenched "fact", presented as an ideal model, that is being overturned by considering related fields of investigation is from the top of the Belaga Formation, outcropping widely across onshore Sarawak. This is considered to be almost phyllitic basement (Liechti, 1960; Petronas, 1999) and is dated as Palaeocene to Late Eocene in age, terminated by the regional Rajang Unconformity; the primary stratigraphic division of the Cenozoic stratigraphic record across western Borneo. Searching out the few old wells offshore Sarawak that drilled into equivalent marine claystone finds data from independent discipline that challenges the concept of metasedimentary basement capped by a major scale break in the geological record. The clays in these wells have sonic log values of around 80µsecs/ft (only moderate compaction), and thermal maturity indicators (spore colour and vitrinite reflectance) that were the same as the overlying Oligocene Cycle I sediments. In 2023 geologists working for Petroleum Sarawak Berhad (Petros) decided that the sonic log and thermal maturity data were reliable and carried out new fieldwork onshore, confirming that the concept of the Rajang Unconformity is a massive over-simplification of a much longer lasting and complex transition, which no longer includes the upper Belaga Formation as economic basement for hydrocarbon exploration. The results of this fieldwork are not yet published, but the key point here is that independent data from non-stratigraphic analyses forced a "*back to the drawing board*" re-examination of a long accepted, idealised concept. Such an advance in science is outside the scope of model-based deductive studies.

The abductive approach is therefore an evidence and test-rich process, which specifically anticipates that the complex and unique history of a tectonically active basin cannot be reliably predicted by an ideal, tectonically passive model. For sedimentary geology there is a fundamental difference in that in tectonically active basins the relative sea-level, and hence sediment accommodation space and sediment supply, do not change in a geographically even fashion, as would be expected from eustatic or epeirogenic sea-level changes (e.g., Matenco and Haq, 2020). Instead, the rapid tilting and movement of both depocenters and hinterland architecture can change in ways that could not be predicted from passive margin type models. Many times, in SE Asia such changes were very rapid and they became the primary tectono-stratigraphic divisions, but they require a specific mode of study.

In contrast to Matenco and Haq (2020) the method proposed here does not seek to define a "*first-principle conceptual model*" but instead it seeks a way to study the predictable properties of the possibly unique tectono-stratigraphic framework. For this the role of Walther's Law is

important, and the derivative that at major unconformities the palaeogeography of the tectonic movement will be related to the contrast in facies across the unconformity.

## COMPONENTS OF THE NEW METHOD

There are three parts to the application of the new method (it is the traditional approach of geologists, slightly updated). The first is quality control on inputs, and the second is the actual shift from model-based to evidence-based processes, and thirdly is the application of a null hypothesis test. This third part does not necessarily come at the end and should also be applied to intermediate interpretations of individual disciplines. In the following paragraphs the reliability of data is discussed, with examples that challenge simple model-based thinking.

The shift in emphasis from model to evidence highlights the importance of questioning input data, current paradigms and the need to repeatedly test ideas. This high level of scepticism contrasts with the deductive, model-based approaches, where it is assumed that both the guiding model is established and trustworthy, and also that a new study will give a result that resembles the ideal model. In such studies it is common practise to accumulate data with some redundancy so that a modal value will appear, and bad data (e.g., biostratigraphy from caved fossils in drill cuttings samples etc.) be deviants on the edge of the bell-curve predicted by the ideal model. The workflow is designed only to deduce from the model, and so if reality deviates from such expectations, it is unlikely to be recognised. For example, biostratigraphy analysis looking for cyclicity reflecting high and low stands in sea-level has a simple format seeking to spot acmes of transgression that are candidates for maximum flooding surfaces. In North Luconia, however, the wells encounter a hemipelagic, globigerina marl up to a hundred metres thick that has the regional Doust MMU at its base. This is a singular flooding event an order of magnitude greater than the anticipated smaller and repeated eustatic events. Not only is this stratigraphic unit then outside the model considered by sequence stratigraphers, but its presence impacts the workflow applied to the thick section below the marl. This abundantly micro-fossiliferous marl facies caves into cuttings samples from the poorly fossiliferous siliciclastic section below, drowning out the very different in-situ fauna. As a result, expensive sidewall core samples are required to carry out palaeontology below the unconformity, and one such sidewall core in this deeper section is more valuable (much higher ranking of reliability) than multiple drill cutting samples from a similar depth. Such ranking of samples for quality is part of the evidence-based method, but it deviates from the deductive workflow that assumes eustasy and passive basin conditions. It can be argued that good workers following eustatic-based workflows can accommodate such local anomalies. However, in doing so they are already moving from simple deduction to abduction, and the best approach is to continue an examination of this major stratigraphic event (Doust MMU) that dwarfs any smaller cycles that the model assumed had controlled stratigraphy. The eustatic model did not predict this major stratigraphic event, so what else might it have missed? To reduce this to a single sentence; model-based deduction is naturally prone to confirmation bias, but abduction is necessarily sceptical and wants to test the empirical framework.

The danger of confirmation bias in using a stratigraphic model with a simple binary

polarity (such as accommodation space either increasing or decreasing), and which specifically assumes a constant location for the hinterland and the depocenter (see Posamentier et al., 1988), has already been highlighted by Miall (1992) and Miall and Miall (2002). The conditions found in an area such as North Luconia indicate that this is a tectonically active basin and calls out for both scepticism and a different approach, using different reasoning, to answer new questions and avoid confirmation bias.

## PRECISION, ACCURACY, AND TRUENESS FROM EXAMPLES

Observations in geology must be ranked for what other sciences present as measurement error bars. However, in geology these are usually two or three independent dimensions of precision, accuracy and trueness (see below). That such basic quality controls are overlooked may surprise other scientists reading geology, but while a so called "hard-science" paper would be rejected for not including observation error-bars, this type of documentation is invariably absent from geology reports. Error bars and precision in naming conventions are discussed in the first essay in this series (Lunt and Luan, 2023), but here some more applied examples are given.

Figure 1 is from a well report in SE Asia just over ten years old and is quite typical. This well drilled a reefal objective, originally deposited close to sea level. Perhaps the most important information from the well samples is the contrast between the reef (mainly its age), and the age of the overlying siliciclastic that contain outer neritic to bathyal microfossils (right-hand columns on Figure 1). There is assumed to be missing stratigraphic section after the termination of the reef, while the reef flanks were being onlapped. Yet the technical report containing the original of Figure 1 misses this, and logs deep marine conditions to total depth as well as the same nannofossil zone. Note that Zone (NN9) lasted almost exactly 1 million years (Gradstein et al., 2012), so a very rapid transition from shallow to deep marine (in less than a million years) is directly implied if the marker fossils in drill cuttings below the top of the limestone are in situ. Clearly there is much clarification needed on the reliability of the analyses that went into this commercial report.

We propose that this lack of scientific discipline and analytical rigour is a product of our over-reliance on models, rather than data, to guide our thinking. The real well example in Figure 1 can be correlated to outcrops onshore and other wells in this play, and the reefal limestone should contain a diverse suit of carbonate facies, larger foraminifera that became extinct at the top of Letter Stage Lower Tf (older than about 12 Ma). No larger foraminifera were looked for in the limestone drilled by the well (three drill cuttings samples examined). Confirmation of a Lower Tf age for this high relief pinnacle, perhaps refined by strontium dating, would have greatly improved our understanding that a reef subsided to deep marine conditions a little before 12 Ma, and then the flanks were onlapped up to the crest all before a time between 10.5 to 9.5 Ma when the NN9 clays covered this location. A simple change to workflow would have resulted in a better description of both very rapid and high magnitude (many hundreds of meters subsidence) tectono-stratigraphic change. Reporting aspects of reliability in this case would consist of highlighting the type of sample examined, and cross-referencing the appropriate, facies-related biostratigraphy and strontium dating,

matching correlatable instances nearby, as well as citing the reliability of fauna indicating outer neritic to bathyal conditions as well as a geohistory plot that would also indicate limits for the magnitude of new accommodation space that had been generated by the reefal subsidence. This mid Middle Miocene reef crest is currently over two kilometres below modern sea-level. Geohistory analysis would begin to quantify how much of this was isostatic response to later sediment loading, and analysis of the siliciclastic fill would help identify if the tectonic component of subsidence was mostly active at the termination of the reefal facies.

## MUNGING DATA

In addition to the evaluation of sample quality, geological context, and general quality control on analytical methods, there is a need to be aware of a related but opposite effect when reading reports. This has the slang name of "munging", also known as data wrangling. This is the process of converting data into another format, often as a precursor to data integration (originating in computer sciences). This is often a desirable and necessary step, such as converting the depths of fossil datums from biostratigraphy in a well to a time scale for use in geohistory or basin modelling. There is a long-recognised risk which reflects the

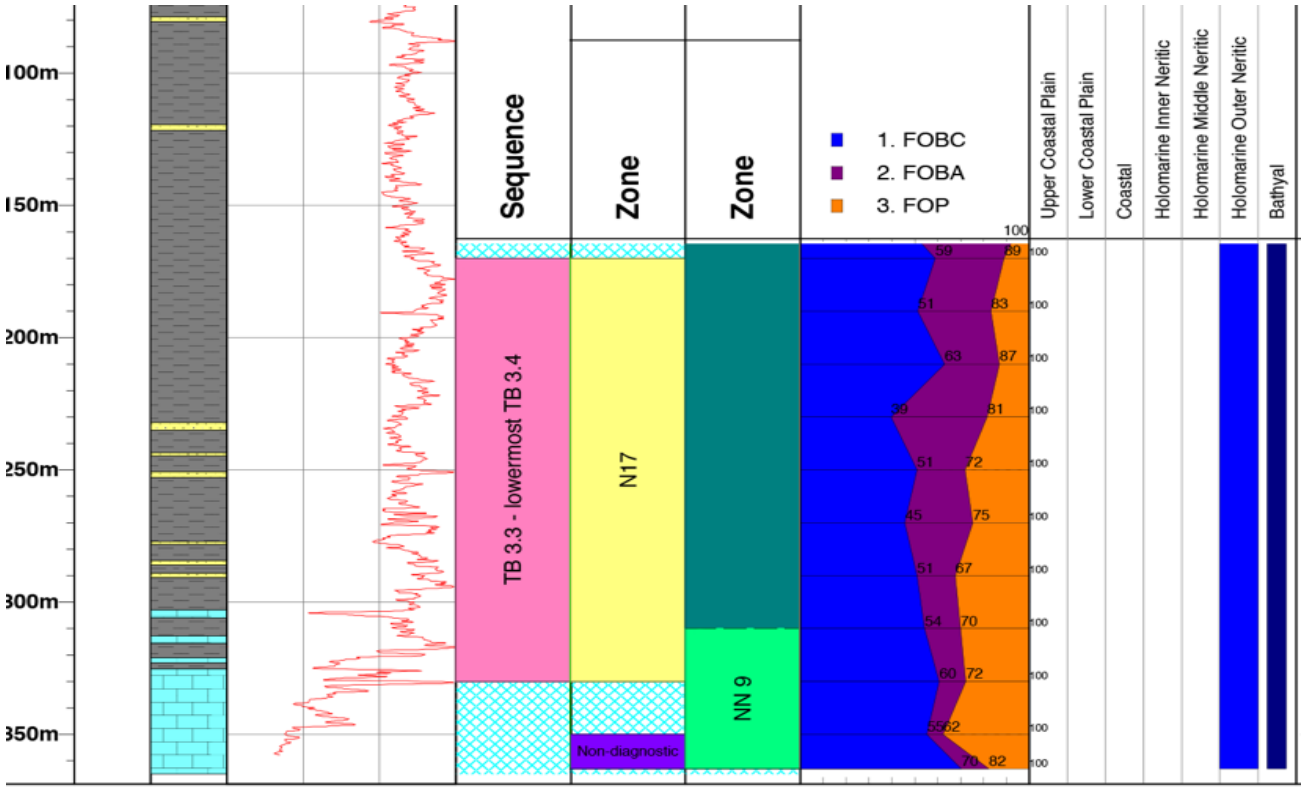


***Figure 1**: A 2008 well (actual report data but cropped to disguise source) showing complete lack of consideration of data reliability. This type of example in is not unique, or even rare. No larger foraminifera were looked for in the limestone. A well a few tens of kilometres away and outcrops onshore have an assumed equivalent limestone with Lower Tf foraminiferal faunas, and these are overlain by only very slightly younger siliciclastic with deep marine microfossils.*

supposed origin of the term; that the name is derived from acronym “Mash Until No Good”. Hence munging is a colloquial term used to flag data that may have been “corrected” prior to input into a database, but in changing it is rendered useless or misleading.

Figure 2 is from a real well in SE Asia. On the right is the column from the final well report with lithostratigraphy and the operator’s report of the environment of deposition in words. On the left is the analysis report, flagging that the work was based on sidewall cores. Examples like this are not uncommon, but this one was chosen because analyses were based on sidewall cores. The operator’s column on the right side first “munges” the uppermost sands into a shallower environment of deposition, just below the modern seafloor of about 60 m water depth, even though these sands are not sub-Recent but below a regional unconformity. That is, the text on the right side contradicts the SWC analyses that directly indicated bathyal “upper slope” or “lower slope” in the middle part of the diagram. The sands in the middle of the well are marked by a question mark, which is still munged data as quite clearly there is no evidence for anything other than very deep marine conditions. As with the example given in previous paragraphs, such poor-quality control is common in both commercial and academic reports, and such misrepresentations are then re-cited and become accepted. This is symptomatic of a migration away from analytical data. It can only be assumed that such workers assume that some larger model will guide interpretation, as they are certainly not looking to the evidence to test geological ideas.

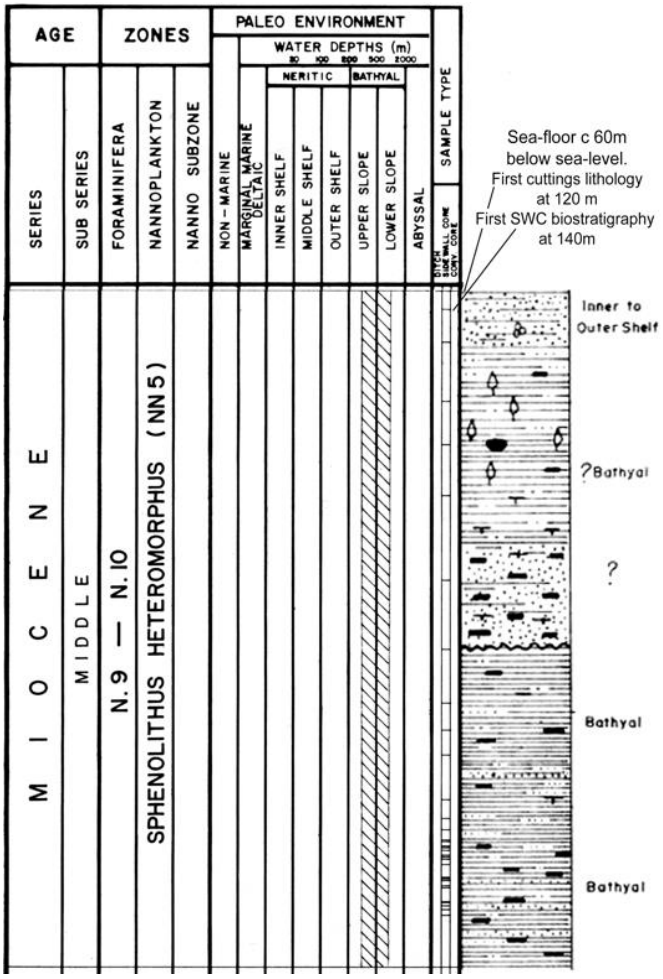


***Figure 2:*** *A disguised real well, with final well report summary column on the right, and the micropalaeontology summary on the left, with a column indicating that analyses were from sidewall cores. The top of the lithology column shown is 60 metres below the modern sea floor.*

## WHEN OLD SCHOOL ANALYSES TRUMP HIGH TECH METHODS

Depending on the type of analysis, the dimensions of precision and accuracy may overlap. However, the following is an example that has measurement precision, distinct from accuracy, as well as trueness. Figure 3 is updated from Luan and Lunt (2022) for strontium isotopic dating in the north Madura area off eastern Java, constrained by biostratigraphy, including the mass extinction of forms at the Eocene-Oligocene boundary that is well

recognised and well dated (33.9 Ma Premoli-Silva & Jenkins, 2006). This could be used as an example of HARKing (Kerr, 1998) as the apparently precise data yielded by the new, and expensive, high-tech method of strontium isotopic dating was inadvertently assumed to be "better" than old-school biostratigraphy, and a new regional stratigraphic correlation was

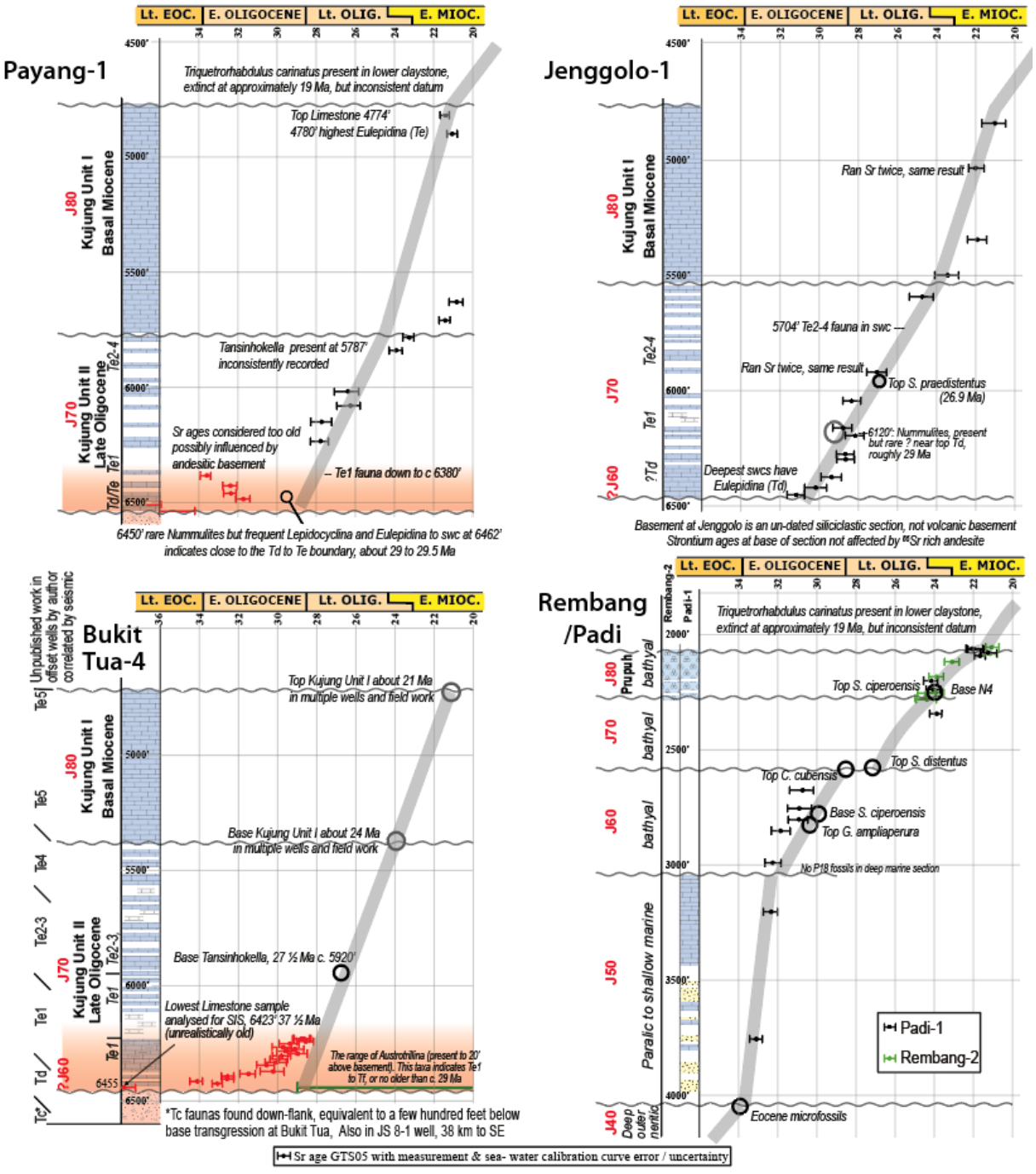


***Figure 3:*** *Summary of strontium dating in North Madura wells (modified from Luan and Lunt 2022). This summary of strontium dating shows the independent dimensions of precision (usually high), accuracy (repeatability; also, apparently high) but low trueness in the wells drilled over andesitic volcanics, indicated by the orange shading. The error bars on the data points are the mass-spectrometer measurement precision.*

integrated to reflect this (Maynard and Morgan, 2005).

Sidewall core and core samples to within a few metres of the basement unconformity contain mid Oligocene foraminifera and lack the Eocene foraminifera predicted by strontium dating in some wells. Bukit Tua-4 has a coherent trend of many Sr ages through the entire Early Oligocene and well into the Eocene that have good machine measurement precision, and both the trend and close spacing of adjacent samples suggests good accuracy. However, from biostratigraphy we know this cannot be true. No other record in the world has species that became extinct at the end-Eocene "Grand Coupure" paleontological event, surviving in some niches for several millions of years into the Oligocene. The Jenggolo-1 shown on the upper right is also from north Madura but located over metasedimentary basement, and this well has a combination of strontium and biostratigraphy data that match each other (and seismic correlation to the palaeontology in the other wells). Lower right is a reference curve for tectono-stratigraphy through the entire Oligocene in eastern Java, where strontium and biostratigraphy again closely agree, although the machine precision of the strontium data is scattered (lower accuracy) and locally outside the measurement range. The biostratigraphy datums marked are biassed to the best age datums.

The reason for failure in this test of trueness in the basal sections in Payang-1 and numerous Bukit Tua wells is thought to be the andesitic basement that is probably a source of $^{86}Sr$. This seems to have permeated a few tens of metres into the overlying sedimentary section, reducing in its effect with distance from the source, and this has artificially changed the $^{87}Sr/^{86}Sr$ ratio to give a falsely old age.

## THE NULL-HYPOTHESIS TEST

The definition in experimental sciences for statistical significance is the arbitrary $P<0.5$ (Fisher, 1925), or a that a result has less than a 1 in 20 chance of being a null value: where null means indistinguishable from just normal, random variation and "back-ground noise". Failure to consider a null-hypothesis test in non-experimental sciences is the equivalent of p-hacking in the experimental sciences (Simonsohn et al., 2013), or HARKing (Kerr, 1998), as all of these are the proposal of an arbitrary result or model, with selected facts being presented as a tested conclusion, when in fact no independent test occurred.

In the non-experimental sciences, which lack both controlled experiments and numerical data for statistics, the null-hypothesis test simply asks if an interpretation would stand out significantly from natural variation in data. This is usually a judgment call, such as agreement that if different workers, faced with the same data and same methods (training in sedimentology, geochemistry etc.), would almost always recognise a similar result. For example, the gamma ray log has long been used to distinguish clays (high gamma ray values) from sands, limestone, and the metal casing in a drill-hole (all low gamma). Before logging-while-drilling was used, the wireline gamma log was run multiple times in a hole section, first as a complete scan, but later as a tool on other logging runs to pick a depth-control point (a casing point or a distinct marker bed) so that all wireline logs were calibrated in depth. Such multiple gamma log runs are often slightly different from each other. This might be from the tool facing different parts of the hole on each run, as well as some background noise and variability. Patterns on one logging run, such as modest fining or coarsening trends of sediment, or individual gamma peaks as

candidate flooding surfaces, might appear to be different on another run over the same section (often the case for small gamma log peaks). What passes the null hypothesis test is the lithostratigraphic pattern that an independent observer can agree as being distinctly outside this range of known observation variability.

This is the reverse of an appeal to authority. That is, once trained, a student would, for the most part, replicate the work of his tutor. In other words, no technique should require the trusted eye of the inventor to see the desired result. The schemes of biostratigraphy, for example, have been successfully transferred to multiple new generations. It is possible that poor training (such as lack of expertise in the specialised larger foraminifera) may produce some variation and degradation, but the evolutionary history of morphotypes considered species has been tested and documented, including the ranking of easier and more difficult to identify markers. This becomes peer review of an analytical method, after which, if an exceptional observation is encountered – such as a gap in ages at a suspected unconformity – it is more likely to be a stratigraphic feature than a null-type error in observation.

Concerns regarding distinction from a null hypothesis occur more in subjects that lack tested frameworks or cross-linking with related analyses (examples of cross-linked analyses include vitrinite reflectance with spore colour and logs measuring compaction or density; foraminiferal zones with nannofossils and Sr dating etc.). For example, while chemostratigraphy has had many successful applications, it rarely has such a unique signature that correlation between sites is incontrovertibly significant. Similarly, interpretation of gravity and magnetic anomalies are non-unique because multiple theoretical solutions are always possible, and distinction from a null interpretation usually relies on integration with geological data.

Without deliberately asking for faith based on authority, there is often subliminal pleading for confidence towards a favoured conclusion because it is “high-tech” or a new, ideal model from a reputable source. For example, the Exxon Cycles scheme was still a new technology when Miall (1992) warned of this potential failure in his paper “*Exxon global cycle chart: An event for every occasion*”, which showed how a simple binary character had a low likelihood of passing a null-hypothesis test. This was further elaborated with an example in Miall and Miall (2004) in a criticism of a correlation between Late Cretaceous sections in the Anglo-Paris Basin and in southeast India. The criticised work specifically set out to demonstrate that sea-level changes are globally synchronous and therefore must be eustatically controlled. The critique by Miall and Miall included the quote by Kuhn (1962) “*Results which confirm already accepted theories are paid attention to, while disconfirming results are ignored. Knowing what results should be expected from research, scientists may be able to devise techniques that obtain them.*”

This last sentence succinctly summarises the null hypothesis as a test to counter confirmation bias. Put another way, it might be possible to see the desired result in your data, and the more famous the expert the greater the confidence in such an interpretation, but if a peer with similar experience can interpret different results that are no less plausible, then the test has been failed. This is supposed to be one of the pillars of the peer review system, but selection pressures often means that data is presented in such a way as to make a single outcome the most obvious, or simply

not mentioning that other equally viable interpretations are possible.

The experimental sciences developed a method to circumvent HARKing, by preregistering hypotheses to be tested prior to the experiment being carried out (see Kupferschmidt, 2018). This kind of rigour is usually not possible in the non-experimental sciences, although the third essay in this series (Lunt and Luan, 2024b), promoting the reformational idea of a paradigm shift, contains such an initial estimate of a new hypothesis for peer review. More often however, the test-rich and evidence-based abductive method enforces a similar rigour to eliminate unwarranted confidence in apparently attractive ideas.

## USING WALTHER'S LAW AS A TEST

The excellent range of depositional facies present, and the tropical microfauna they contain, gives SE Asia an additional test of stratigraphic validity through Walther's Law. Perhaps the simplest example of this are the multiple breakup-like subsidence unconformities found at various times around SE Asia. In the sedimentary record there is a fundamental relationship we can exploit. For basins with thick, continuous deposition of a single facies or closely related facies groups (e.g. lower coastal plain to occasionally inner neritic), and then a step-like change to a much deeper – but often also rapidly deposited – bathyal claystone, such as the Agam or Ranong Formations abruptly overlain by Pirak or Yala Formations across north Sumatra and West Thailand area, then this vertical, step-like sedimentary record must reflect a step-like, sudden shift in sedimentary palaeogeography. This is Walther's Law, and the step-like facies shift has been named a non-Waltherian contact (NWC; Lunt, 2019a, b). At the event the entire sedimentary system withdrew, and it cannot leave a portion of a delta behind on some more slowly subsiding, isolated, submarine high. Similarly, any clays within the underlying Agam Formation, deposited before this event, cannot be classified as belonging to the Pirak Formation, and thereby drawn as an interdigitation of the two named lithofacies. Yet both in this basin and elsewhere such misinterpretations are frequently found in reports and publications. The presence of such a non-Waltherian contact in the four dimensions of space and time is a test of correlations, stratigraphic models and basin history. Such tests are significant in that if the facies data is correct, then biostratigraphic or seismic correlation cannot cross an NWC. Being based on a Law and not a theory, this is not an issue of interpretation where workers can agree to disagree.

## THE ABDUCTIVE METHOD SUMMARISED

Abduction is the process of taking different inductively derived hypotheses, applying the data quality controls described above, then testing these hypotheses against each other and any existing geological framework. Continuing the example of the previous paragraph, a non-Waltherian contact can be dated with foraminiferal data, cross-checked by nannofossil and strontium data with the knowledge that no biostratigraphic correlation datum can cross the NWC, and that the NWC must have a facies shift reflected both in micropalaeontology and lithofacies (and be geographically variable in a mappable, meaningful pattern). All these independent disciplines, plus timing of fault activity, can be combined into a single historical framework. This, of course, was an

established method of investigation until the key science of stratigraphy (which links facies in space and time) moved towards an inappropriate ideal model, and for many basins around SE Asia we lost track of the role of tectonic activity and its geographic variability.

The power of abduction is that it can study unique geological scenarios that are well outside ideal models. This might include drilling a feature that was predicted to be a reef but was then found to be a buried volcano. Simple induction changes from "*all pinnacle highs are reefs*" by adding the caveat *"… and some are volcanos*". That is, there is a numerical probability based on ratios of known data that any future well would find either a reef or a volcano. In contrast, abduction proposes hypotheses that volcanos should perhaps have a geographic trend, be of a certain age, be associated with feldspar-rich or tuffaceous sands of a certain age, that uplift during intrusion created shallowing around the flanks (associated with increased feldspar / tuffaceous sands). Abduction is the active leading or forcing of alternative hypotheses through a process of investigation from multiple independent disciplines.

In the case above, it is of course expected that researchers would naturally begin to ask about igneous activity after the first volcano was drilled, in order to prevent another commercial failure, and not just rely on an over-simple percentage risk. But in stratigraphy, with poor quality-controlled data such as in Figure 1 (a real well from an area with volcanic activity just older than the reefs), such critical questions are often not asked. In this example why, within a very short period of time, did such a series of reefs subside to very deep marine conditions? Deductive methods such as eustatic sea-level models cannot begin to tackle such unique questions. As pointed out by Lunt and Luan (2022) simply changing from a eustatic sea-level curve to a local relative sea-level curve fails to honour the scientific principles of both a model-based deduction and an evidence-based local "model". There are many widely cited papers that attempt to carry out such a compromise (e.g. Wong, 1993, Figure 22). Instead, workers must keep separate the model-based concepts while an evidence-based framework is established. The model-based concepts have assumptions that should not be accidentally applied to the empirical framework, such as geographically eustatic sea-level changes does not mean that tectono-stratigraphic changes to sedimentary siliciclastic are also evenly expressed across the basin (e.g. Wong, 1993, Figure 22). It is arguable that the geographic variability of tectono-stratigraphic sequence boundaries is their most important property. Continuing the example shown in Figure 1, mapping out this variability in the event that terminated a series of reefs would include a study of the outcrops located onshore from this well, which have not been visited by professional geologists since the mid 1960s. There it can be seen that reefs grew above sub-aerial volcanic beds, and the carbonates first slowly subsided to form pinnacle reefs, then rapidly subsided and drowned. The age of this rapid subsidence of multiple reefs appears to correlate with a regional subsidence event that is observed a few hundred kilometres further west, which was dated from biostratigraphy in the 1980s. However, in recent years a different age was mistakenly assigned to that western subsidence after miscorrelation with a eustatic sea-level event. That is, a failure to keep model-based assumptions separate from empirical data. The abductive process demands clarification of many geological topics, and ideally confirmatory new analyses. In contrast, the model-based

deductive process survives, and even thrives, with less geological investigation.

This example of buried volcanos near reefs could be argued to be an exception (it was used here only as it followed on from the example chosen to illustrate the point in Figure 1), but the present authors have multiple papers across SE Asia with other exceptions to accepted, over-simplified, but regularly re-cited geology. Each is unique, as SE Asia is highly varied and complex, but all such examples have roots in data quality control and a paucity of independent tests. Rather than give multiple different examples of abduction it might be more beneficial to highlight two roles of abduction in geological studies in tectonically active basins; the first is more obvious and of less value, the second is harder to explain to non-geologists but it is where the greatest values lie.

First, abduction forces a greater degree of resolution and precision on the geological framework. Lunt (2021) demonstrated that an unconformity long considered base Cycle V in offshore west Sarawak was, in fact, overlain by Cycle IV limestones in several wells and therefore the unconformity was the base of Cycle IV (the Doust MMU), and this matched a simpler regional tectono-stratigraphic model. How important is that correction for the seismic pick for base Cycle V to base Cycle IV? A blunt answer is that nobody changes their opinion of an exploration prospect because of a revision and improvement to the regional geological framework. Seismic-dominated studies can be argued to satisfy the Pareto Principle of economics that implies roughly 80% of consequences come from 20% of causes, and that 80% of subsurface understanding can come from the single discipline of seismic interpretation, with the more abstract field of stratigraphic geology adding relatively minor commercial value. (This simple logic is slightly undermined by the relative costs of expensive seismic acquisition and processing, versus cheap well analytical studies.)

The second effect of abduction is the reverse of the 80:20 rule; that by forcing resolution of the geological framework an unanticipated divergence may appear, and we are forced to re-invent the geology. In such settings the low cost, abductive, analytical geology delivers 80% of the value, in an innovative paradigm shift that can lead to whole new geological concepts or exploration plays. Of course, this does not happen very often, but it is most likely after a prolonged period of stagnation. Details of some new concepts in SE Asian geology are presented in other essays in this series (Lunt and Luan, 2023; 2024b) as they are best discussed along with comments on data reliability. A short example of the principle is the mid Oligocene tectono-stratigraphic event in eastern Java, noted by Matthews and Bransden (1995) to have both localised uplift and subsidence visible on seismic. This movement was focused along a narrow geographic trend in the zone south of the Rembang Line, along almost all of Java to the Madura Straits (Luan and Lunt, 2021) where entire proto-petroleum systems were locally uplifted, slightly eroded and the tops of the structural highs then developed reefs during Late Oligocene through Early Miocene subsidence. Seismic data below this carbonate is poor but later Neogene burial led to source rock maturation, and sub-Kujung Limestone structural traps with sands are possible but have only once been deliberately drilled. This is the reverse of the established concept of Kujung reefs growing on ancient basement highs. It is also inconsistent with the concept of the trough under Java being an extensional back-arc basin.

Such changes to basic geology are not always going to yield new hydrocarbon prospects but they can radically change key exploration risks. For example, the Stage III (Early Miocene to early Middle Miocene) of west Sabah is on record as being deformed economic basement (Petronas, 1999), as is the Belaga Formation of Sarawak (see introductory section), but both these sedimentary sections are misunderstood due to over-simple descriptions of their bounding unconformities. Both have organic rich facies, locally oil seeps (in mis-dated outcrops), and in certain areas they retained low levels of thermal maturity until later Miocene burial.

Only evidence-based, abductive, geology can reconstruct such unique problems, and by definition this will lead to either new or strongly re-risked plays that are "*off the creaming curve*". This is oil industry slang for novel discoveries with potential for high value, as distinct from ever-smaller replicas of known hydrocarbon plays. In 1982 Michael Halbouty edited an AAPG memoir which outlined the importance of developing new concepts to find the next generation of hydrocarbons. However, if we cannot study, measure and predict the tectono-stratigraphic complexity of SE Asia sedimentary basins we cannot begin such a deliberate search. Modern seismic data is excellent, but in the decades taken to develop the acquisition, processing and display tools for this data we have neglected the natural geological framework in favour of a pastiche of science based on passive margin models that simply fail (do not match observed data) in many areas with oil seeps and unexplored sediments.

As noted above, this is not a wholly new method but rather a return to classical methods. The prose of van Bemmelen (1949) is rich in abductive thinking. It weighs up primitive sets of gravity and volcanic data with rudimentary stratigraphy and structural ideas (precision, accuracy and trueness of each observation). The Geological Survey in Malaysia was equally active until the mid 1970s with work such as Liechti (1960) through to Leong (1974), the latter author going into detail on why the Ayer and Kuamut Mélanges of east Sabah are extensional and not compressional deposits. However, since then an aggregation mentality has dominated geology – which assumes as more data is gathered, possibly in small increments, it will inevitably trend towards better conclusions. This is not the case as the process requires that specific hypothesis must be constructed, tested and re-formulated in a deliberate fashion. To compound this problem, academia and industry have moved towards the simpler, cheaper, deductive processes. For example, for thirty years the industry assumed it could deduce stratigraphy from a eustatic, or modified eustatic model, even though the largest candidate events such as the mid-Oligocene sea-level fall and near base Late Miocene sea-level fall have never been identified in the region (Saller et al., 1993, Matthews and Bransden 1995; Lunt, 2014; Morley et al., 2021).

Such prolonged stagnation in critical thinking (lacking repeated, independent testing of different observations), while old concepts are re-cited and become entrenched in our mental models, forms a barrier to the adoption of new processes. This is because a return to abductive critical thinking will first break-down familiar old concepts before re-building a new framework. To take only small steps along this path is a destructive process that will probably meet resistance in peer review, until it is widely recognised that a paradigm shift is in process. A following paper examines several features around

Sundaland to illustrate that that not only is a paradigm shift possible but that it well underway.

## CONCLUSIONS

There is no reason to expect that the replication crisis in the experimental sciences does not also affect geology, where it might be better termed a validation crisis. In the non-experimental sciences the model-based approach mostly escapes testing because the source of deduction is an ideal model, which is assumed to be correct. In the earth sciences, especially in tectonically active basins where each area has a unique stratigraphic history, the only way to avoid a validation crisis is to define a framework (a documented geological account, not a conceptual model) and then seek to test it. What appears to be happening in SE Asia, both from onshore geological surveys, and offshore from oil companies, is that fewer and fewer analyses or studies capable of testing a geological history are being carried out. Such analyses, quality controlled and documented accordingly, are essential for independent tests of models and interpretations. Miall (1992) showed the great difficulty of refuting an ideal model even with good age data for the events. So the trend to decreasing the amount of independent age, and facies data exponentially reduces our ability to test basic assumptions, and the science stagnates. We must also move documentation away from simple narratives to carefully constructed technical accounts that includes notes on the reliability of input data, the tests applied, and distinction of results from a null hypothesis.

This review emphasises that ideal models are very poor analogues from which to develop an understanding of complex and unique geological scenarios. They may be an unavoidable starting point where there is no other data, but the acquisition of diverse data is paramount, and an abductive approach, of developing several methods of investigation and cross-checking each of these, to refute some of the alternative working hypotheses, is the only way to progress. Where tests have been carried out, several long-established and widely cited models in SE Asian geology have been found to fail (e.g. Lunt, 2022 on subduction of a proto-South China Sea plate), yet we lack an objective process to construct the replacement. This review discusses some ideas to develop a replacement process.

## REFERENCES


Baker, M., 2016. "1,500 scientists lift the lid on reproducibility". Nature (News Feature). Springer Nature. 533 (7604): 452–454. *DOI: 10.1038/533452a*

van Bemmelen, R.W., 1949. The Geology of Indonesia. The Hague, Netherlands: Government Printing Office., 997

Fisher, R.A. 1925. Statistical Methods for Research Workers. Oliver and Boyd, Edinburgh, 237 pp.

Gradstein, F.M., Ogg, J.G., Schmitz, M.D., and Ogg, G.M., 2012. The Geological Time Scale 2012 (1). Elsevier, Amsterdam., 1144

Halbouty, M.T., 1982. The Deliberate Search for the Subtle Trap., DOI: 10.1306/M32427

Haq, B.U., Hardenbol, P.R., and Vail, P.R., 1987. Chronology of fluctuating sea levels since the Triassic. Science 235, 1156-1167

Haq, B.U., Hardenbol, J., and Vail, P.R., 1988. Mesozoic and Cenozoic chronostratigraphy and eustatic cycles. Sea-level changes; An integrated approach SEPM Special Publication 42, 71-108

Ioannidis, J.P.A., 2005. Why most published research findings are false. PloS Med 2(8),

Kerr, N. L., 1998. "HARKing: Hypothesizing after the results are known". Personality and Social Psychology Review. 2 (3): 196–217.
*DOI:10.1207/s15327957pspr0203_4*

Kuhn, T.S., 1962. The structure of scientific revolutions. Chicago & London: Phoenix Books, The University of Chicago Press., 196

Kupferschmidt, K., 2018. More and more scientists are preregistering their studies. Should you? Science (News), 18 Sept. 2018.
*DOI: 10.1126/science.aav4786*
*https://www.science.org/content/article/more-and-more-scientists-are-preregistering-their-studies-should-you*

Leong K.M., 1974. The Geology and Mineral Resources of Upper Segama Valley and Darvel Bay, Sabah (4). Malaysia Geological Survey, Borneo Region., 354

Liechti, P., 1960. The geology of Sarawak, Brunei and the western part of North Borneo (Bulletin 3). Kuching Sarawak: Geological Survey Department British Territories in Borneo.

Luan, X., and Lunt, P., 2021. Latest Eocene and Oligocene tectonic controls on carbonate deposition in eastern Java and the south Makassar Straits, Indonesia. Journal of Asian Earth Sciences 220, 104900
*DOI: 10.1016/j.jseaes.2021.104900*

Luan, X., and Lunt, P., 2022. Controls on Early Miocene carbonate and siliciclastic deposition in eastern Java and south Makassar Straits, Indonesia. Journal of Asian Earth Sciences 227, 105091
*DOI: 10.1016/j.jseaes.2022.105091*

Lunt, P., 2014. A review of the foraminiferal biostratigraphy of the Melinau Limestone, Sarawak. Berita Sedimentologi 29, 41-52

Lunt, P., 2019a. A new view of integrating stratigraphic and tectonic analysis in South China Sea and north Borneo basins. Journal of Asian Earth Sciences 177, 220-239

Lunt, P., 2019b. Partitioned transtensional Cenozoic stratigraphic development of North Sumatra. Marine and Petroleum Geology 106, 1-16

Lunt, P., 2021. Tectono-stratigraphic framework of Luconia carbonates. SEPM Special publication, Cenozoic Isolated Carbonate Platforms – Focus Southeast Asia 114,
*DOI: 10.2110/sepmsp.114.08*

Lunt, P., 2022. Re-examination of the Base Miocene Unconformity in West Sabah and stratigraphic evidence against a slab-pull subduction model. Journal of Asian Earth Sciences 230, 105193
*DOI: 10.1016/j.jseaes.2022.105193*

Lunt, P. and Luan, X., 2023. Giving names to features in geology; the choice between subjective listing or researching objective natural divisions. Berita Sedimentologi, 49 (2), 28 – 55.

Lunt, P. and Luan, X., 2024b. When, how much, how fast, and why it matters. A quantitative view of stratigraphy and the emergence of a new paradigm. Berita Sedimentologi, 49 (3), 18 - 46.

Matenco, L.C., and Haq, B.U., 2020. Multi-scale depositional successions in tectonic settings. Earth-Science Reviews 200, 102991.
*DOI: 10.1016/j.earscirev.2019.102991*

Matthews, S.J., and Bransden, P.J.E., 1995. Late Cretaceous and Cenozoic tectono-stratigraphic developments of the East Java Sea Basin, Indonesia. Marine and Petroleum Geology 12(5), 499-510.

Maynard, K., and Morgan, W.A., 2005. Appraisal of a complex, platform carbonate, Bukit Tua discovery, Ketapang PSC, East Java Basin, Indonesia. Proceedings from Proceedings Indonesian Petroleum Association Convention: 30, 317-330.

Miall, A.D., 1992. Exxon global cycle chart: An event for every occasion. Geology 20, 787-790.

Miall, A.D., 2004. Empiricism and model building in stratigraphy: The historical roots of present-day practices. Stratigraphy 1(1), 3-25.

Miall, C.E., and Miall, A.D., 2002. The Exxon factor: The Roles of Corporate and Academic Science in the Emergence and Legitimation of a New Global Model of Sequence stratigraphy. The Sociological Quarterly 43(2), 307-334.

Miall, A.D., and Miall, C.E., 2004. Empiricism and model-building in stratigraphy: Around the hermeneutic circle in the pursuit of stratigraphic correlation. Stratigraphy 1(1), 22-46.

Morley, R.J., Hasan, S.S., Morley, H.P., Jais, J.H.M., Mansor, A., Aripin, M.R., Nordin, M.H., and Rohaizar, M.H., 2021. Sequence biostratigraphic framework for the Oligocene to Pliocene of Malaysia: High-frequency depositional cycles driven by polar glaciation. Palaeogeography, Palaeoclimatology, Palaeoecology 561, 110058
*DOI: 10.1016/j.palaeo.2020.110058*

Petronas., 1999. The petroleum geology and resources of Malaysia. Kuala Lumpur: Petroliam Nasional Berhad (Petronas)., 665 pp.

Posamentier, H.W., Jervey, M.T., and Vail, P.R., 1988. Eustatic controls on clastic deposition; I-conceptual framework. Sea-level changes; An integrated approach SEPM Special Publication 42, 109-124.
*DOI: 10.2110/pec.88.01.0109*

Premoli-Silva, I., and Jenkins, D.G., 2001. Decision on the Eocene-Oligocene boundary stratotype. Episodes 16(3), 379-382.

Saller, A., Armin, R., Ichram, L.O., and Glenn-Sullivan, C., 1993. Sequence stratigraphy of aggrading and backstepping carbonate shelves, Oligocene, Central Kalimantan, Indonesia. In: R. G. Loucks & J. F. Sarg (Eds.), Carbonate sequence stratigraphy: recent developments and applications. AAPG Memoir 57, 267-290.

Simonsohn, U., Nelson, L.D., and Simmons, J.P., 2013. P-Curve: A Key to the File Drawer. Journal of Experimental Psychology: General, Forthcoming, 14 p.

Wong, R.H.F., 1993. Sequence stratigraphy of the Middle Miocene–Pliocene southern offshore Sandakan, East Sabah. Geological Society Malaysia Bulletin 33, 129-142.