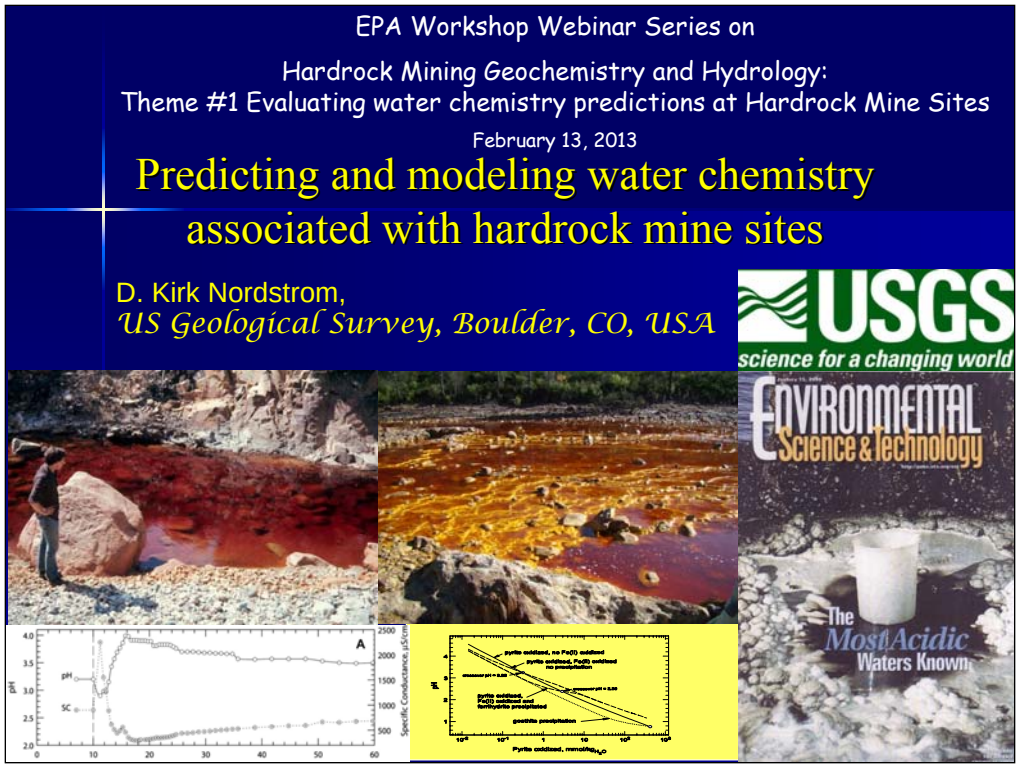
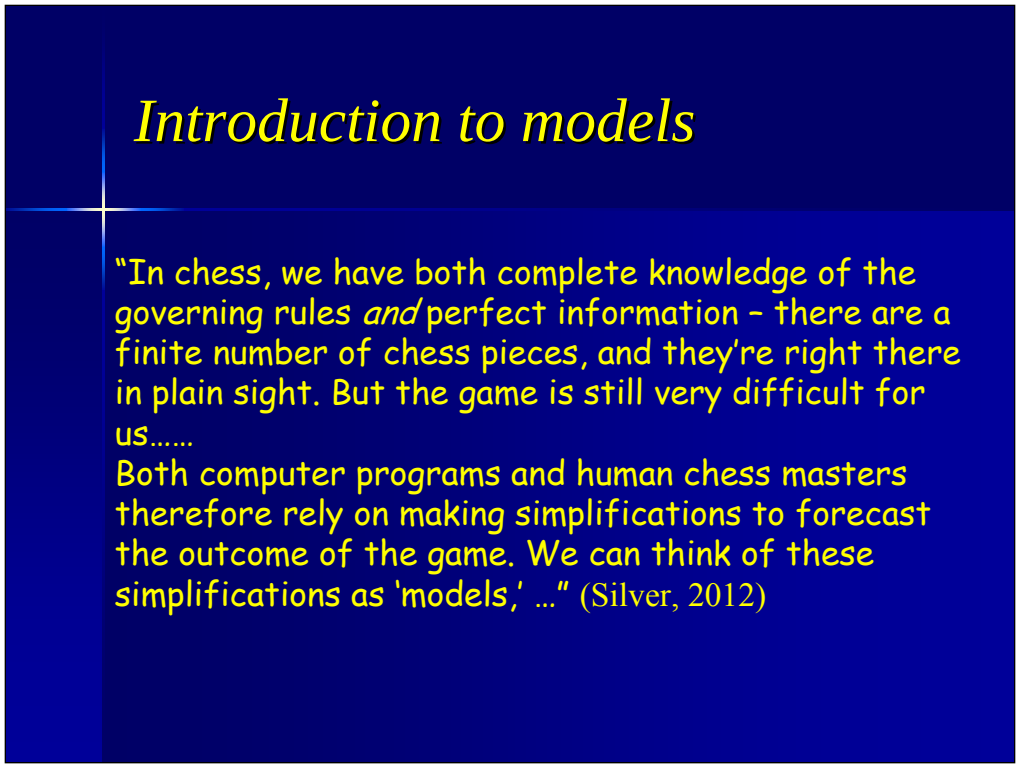
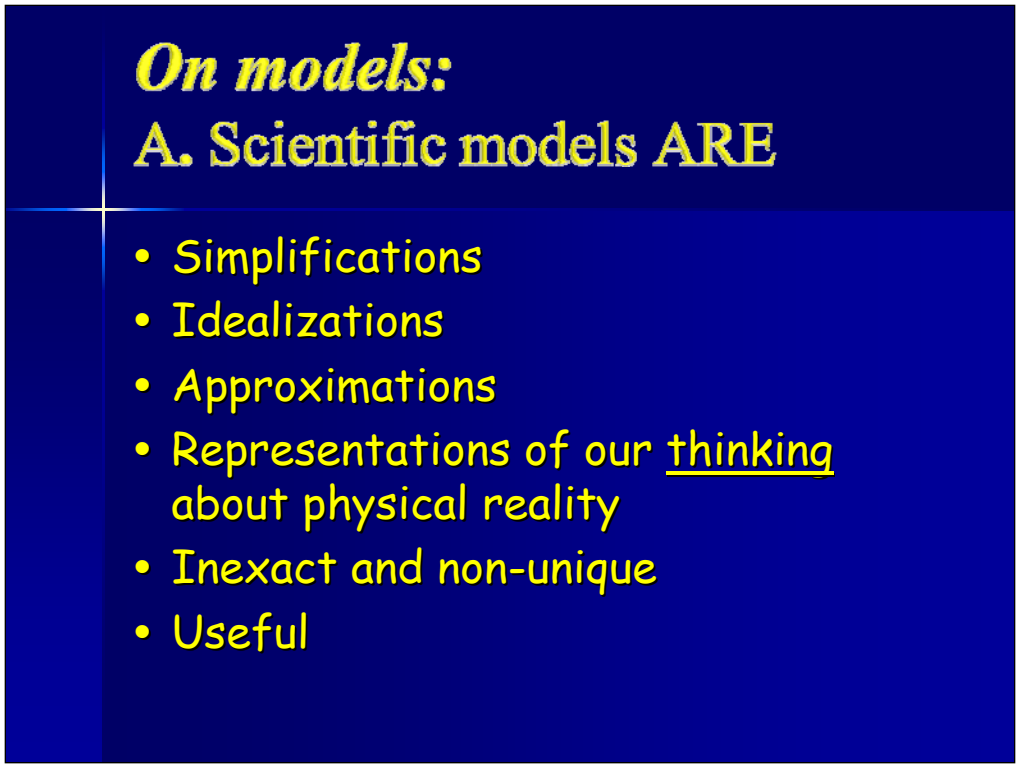


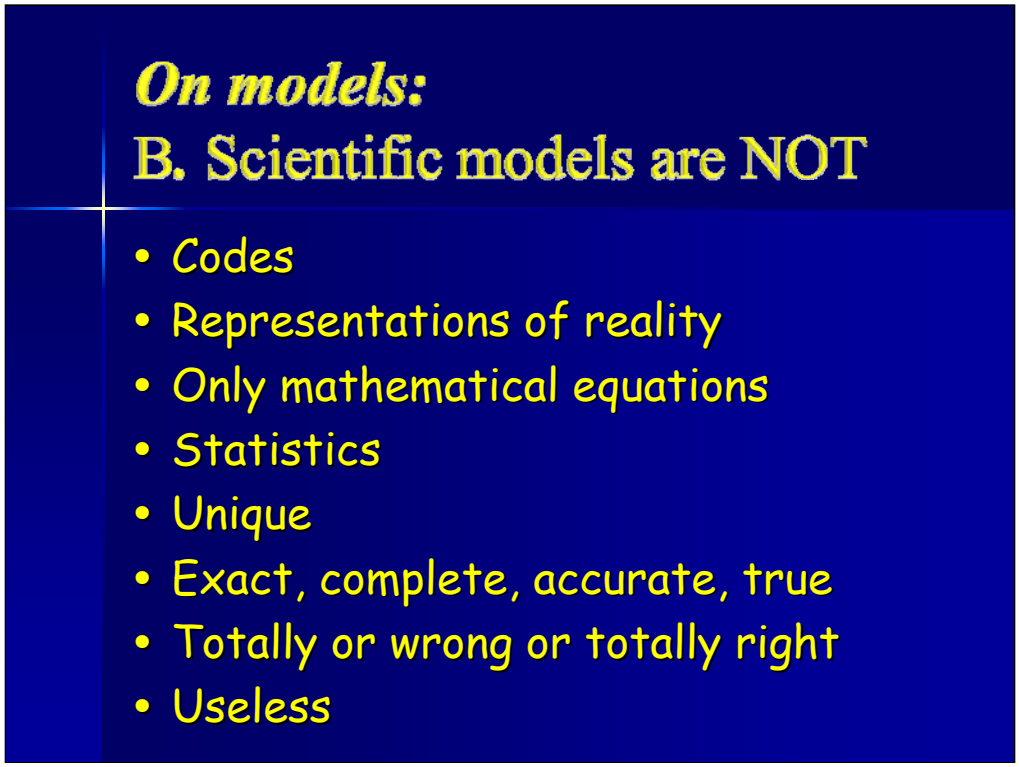
Welcome, everyone, to this webinar. I shall start with a warning that this subject is normally taught over the space of weeks to months to years depending on one's background level of expertise; it is complex and highly technical material. Hence, I'll be presenting a short overview, emphasizing what I consider to be some of the most important aspects within a regulatory framework and some glimpses into the state-of-the-art.



I thought this was a rather good analogy that puts the issue of modeling and its reliability into perspective. Here you have an example of the difficulty of modeling some aspect of the environment. It's worse than chess. Anyone who thinks that a model can provide accurate and reliable knowledge, whether it be characterization or prediction, has not heard of Murphy's Laws. Note that I partly disagree with nate Silver's statement: computer programs don't rely on simplifications. It is the model they embody that relies on simplifications.



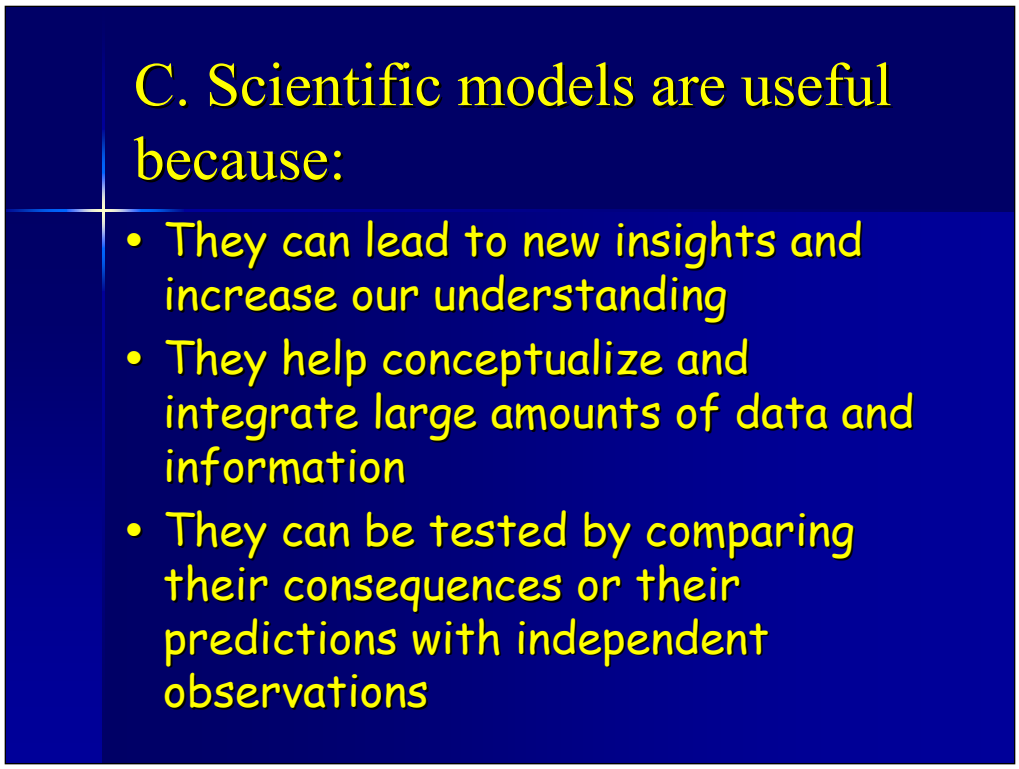
There persists some confusion about what a model is or isn't and it is important to recognize these characteristics. A model is always a simplification. We never know enough for it to be anything else. We simplify by idealizing, hence models are also idealizations and approximations. The important question is: are the approximate results calculated by the model useful? We often say that models represent reality but they don't; they represent our thinking about reality. So the question becomes: How good is our thinking? The answer to that depends on education, training, experience, and creativity. Models come from concepts; if we don't have the concepts right then the models will be flawed. Another consequence of simplification is that models are inexact and non-unique. Nevertheless, models can be very useful – mostly for enhancing our understanding, not necessarily for regulatory purposes. If the regulatory purpose includes improved understanding of the geochemical processes at a mine site, then models can be useful; if regulatory purpose is ONLY focused on compliance requirements or permitting, then models are probably not useful.



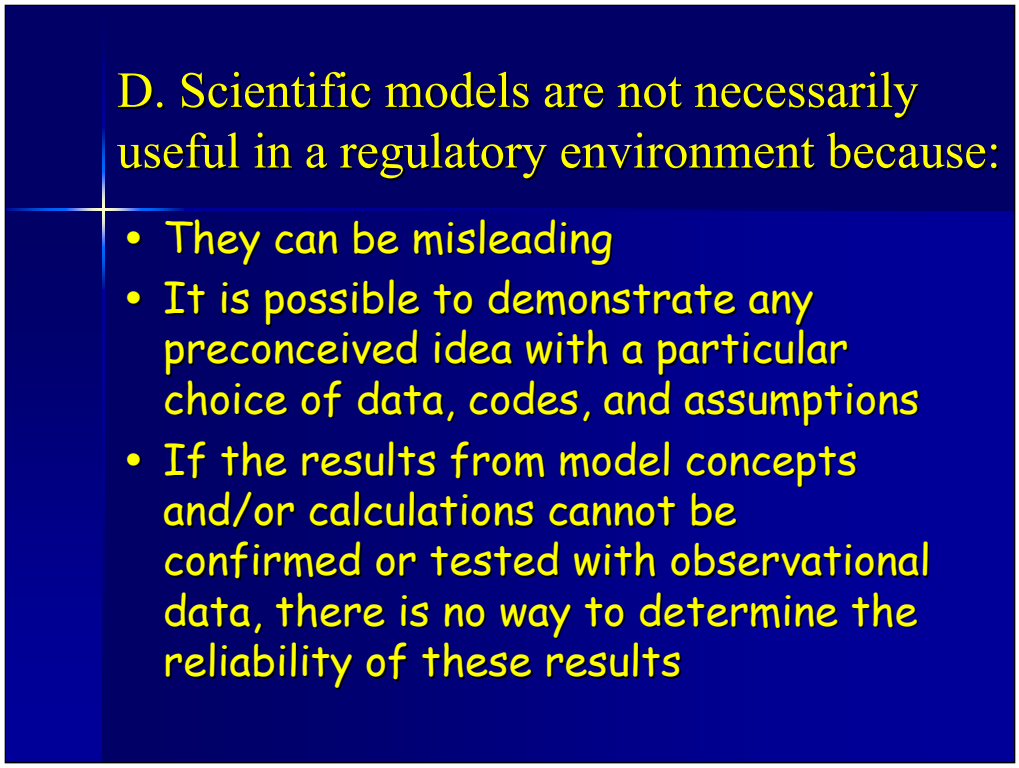
People often refer to the MINTEQ model, or the PHREEQC model, but that is incorrect. These are not models! They are computer programs or codes. If someone says they used the PHREEQC model to compute something, ask them what the model is and remind them that PHREEQC is a code that has gone through numerous versions and incorporates several models (such as the ion-association model and the Pitzer ion-interaction model) and has several databases. They need to spell out which models and databases they are using, not only which codes. Models don't change nearly as much as codes do.

Models are not "representations of reality" because (1) we don't know what reality is to begin with (if we did we wouldn't need a model) and (2) we cannot represent reality, we can only represent our thinking about reality. We don't know what it means to represent reality. We don't know what it means to represent something we can't define and by using language we limit our ability to express that representation. [Gregory quote, p. 1912 of Nordstrom, 2012]. "The minute we begin to talk about this world, however, it somehow becomes transformed into another world, an interpreted world, a world delimited by language." Some people are so immersed in mathematics that they think models are only mathematical equations. Most field-based scientists understand that there is far more to modeling than the application of mathematics. Likewise with statistics. Applying statistics to a set of data does not a scientific model make. It requires interpretation.

Models are not unique, exact, complete, accurate, or true. Models are also not totally wrong nor useless nor totally incorrect. You have to be careful of this tendency to put things into B&W boxes. The world is grey. Some people are fond of the quote "All models are wrong, some are just more useful than others." I don't agree because this statement is another B&W type statement. We shouldn't say that all models are wrong or any model is right because it is too simple a statement. We should simply say that all models are approximations and some are better approximations than others depending on the objectives, the system being studied, and the limitations of the model for the specified conditions.



It is important to recognize that we compare the consequences of our models with independent observations, not the model itself. Einstein and Infeld (1938, The Evolution of Physics) made it clear that we cannot compare our theories with the real world; we can only compare the predictions from our theories with our theory-laden observations of the world.



So if someone says that they have predicted the water chemistry of a pit lake 50 years into the future – the important question to ask is where is the data to show that such a prediction has been tested and shown to have agreed with observation?! How well do these types of models really predict? They would have a hard time answering that question because of the lack of data. If we don't have a confirmation from a test of the model, then we have no basis to have any confidence in the model. Without confirmation, it's guesswork, not science.

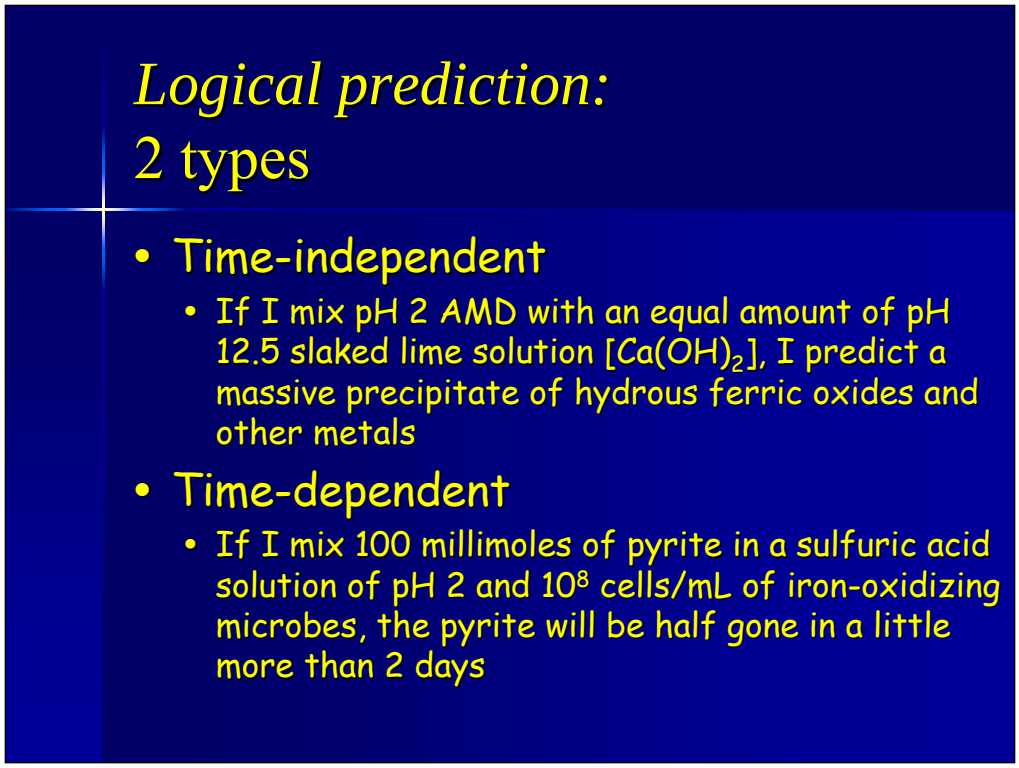
E. Scientific models are not necessarily useful in a regulatory environment because of:

- *The complexity paradox*
  - The more sophisticated a model and the more complex the code, the more difficult it is to test the code and determine if it is working properly, or even to understand how it works [Oreskes, 2000]
- *Loss of meaning and representation*
  - "Needlessly complicated models may fit the noise in a problem rather than the signal, doing a poor job of replicating its underlying structure and causing predictions to be worse." [Silver, 2012]

On predictions

- 2 meanings
  - Logical (or phenomenological) prediction
  - Temporal (or chronological) prediction

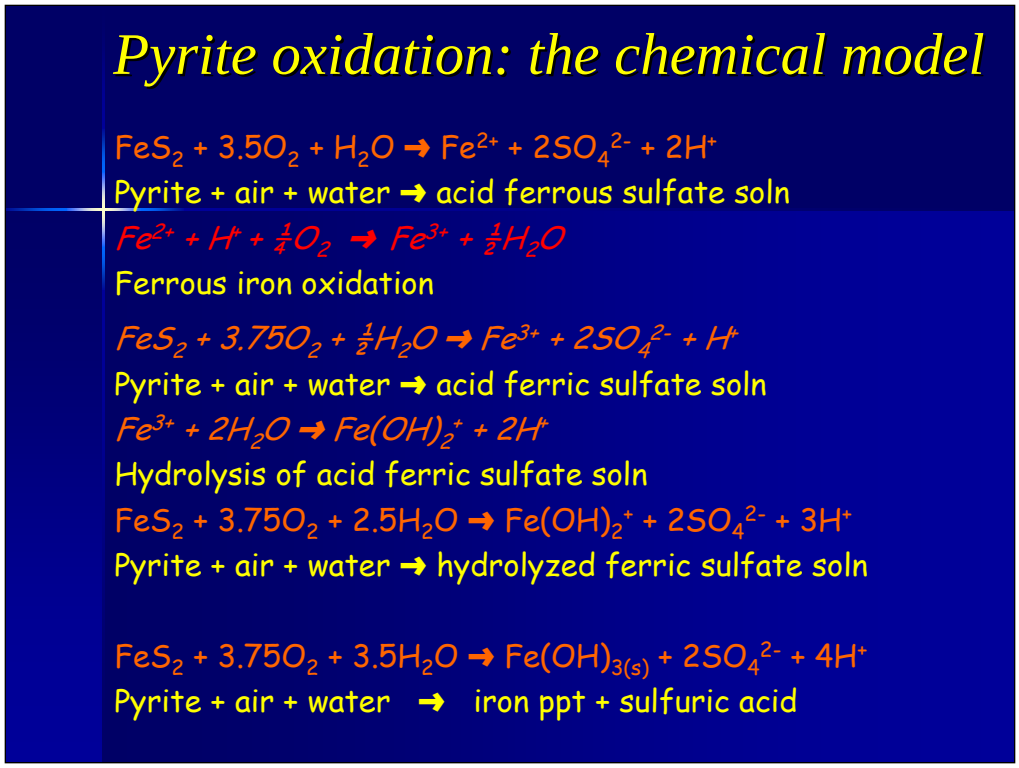
- **Logical prediction:** a prediction based on scientific principles along with necessary assumptions to form a logical construct with testable consequences (what science does)
- **Temporal prediction:** a prediction that foretells the future (betting on horses, predicting the world apocalypse, foretelling the day and hour you will die, etc.; not what science does)



Another source of confusion, even for modelers, is the difference between time-dependent prediction and time-independent prediction. [*read slide here*]. Groundwater modelers and reactive-transport modelers are always working with time as an explicit variable, geochemists often work with time as an implicit or non-existent variable. Geochemists who produce kinetic data on the dissolution or precipitation rates of minerals are working explicitly in time but it is "lab" time which does not necessarily have anything to do with "field" time. In making future predictions, such as predicting groundwater conditions given certain properties and boundary conditions, we are actually trying to make our logical predictions conform to temporal predictions. In so doing, we have to acknowledge that there are some serious limitations in this endeavor because there are some factors that are beyond our ability to predict. This issue was brought out nicely by research in the field of chaos theory and non-linear dynamics (remember the "butterfly effect").

More definitions

- *Chemical model – a theoretical construct that permits the calculation of thermodynamic, kinetic, or quantum mechanical properties of a system*
- *Geochemical model – a chemical model applied to a geologic system*



Pyrite oxidation is actually a rather complex process. It involves the transfer of 14 electrons from the S_2^{2-} entity in pyrite to sulfate. This does not happen in one step. I have not shown all the steps because it is beyond the scope of today's presentation and today's presentation is complicated enough. Also, pyrite is directly oxidized by Fe^{3+} (ferric iron) not oxygen. The overall process comes out to the same results whichever oxidant is used. I have also not mentioned the very important catalysts in the system, iron- and sulfur-oxidizing microbes. That would be another subject for discussion in another workshop. Basically, without the chemoautotrophic bacteria and archaea, the reactions that produce AMD would be much slower.

Example 1: Can we predict water chemistry from pyrite oxidation?

- Yes, BUT ONLY IF
  - We know how much pyrite has oxidized
  - We assume an unlimited supply of O_2
  - We assume equilibrium solution speciation
  - We only consider initial/final states, not intermediate states that require knowledge of reaction rates
  - We assume no other minerals are reacting
- Are these assumptions technically correct?
  - No – but a few waters do approximate these conditions
- We need actual water chemistry data with mass balances to know the amount of pyrite oxidized

Example 1.1 Mass Balance Modeling
Pyrite oxidation with gypsum dissolution

- If a water analysis contains 480 mg/L SO_4 (5 mmol) and mass balances show that 75% came from pyrite oxidation and 25% from gypsum dissolution, then
- 1.875 mmol of FeS_2 dissolved and
- 1.25 mmol of $CaSO_4 \cdot 2H_2O$ dissolved

Here is what we mean by mass balances.

Example 1.2 Mass balance on a natural acidic drainage water [rock is andesite and rhyolite mineralized with pyrite, gypsum, sulfides, etc.]

Solute	Concentration
pH	2.98
Ca	8.72 mM
Mg	4.64 mM
Na	0.35 mM
K	0.019 mM
SO_4	21.1 mM
F	0.40 mM
SiO_2	1.23 mM
Al	3.39 mM
Fe	1.16 mM
Zn	7.63 mM
Cu	0.029 mM

Initial state = pure H_2O/ final state = water composition

Pyrite . . . . . 8.66 mmol/kg
Gypsum . . . . . 3.78 mmol/kg
Dolomite . . . . . 4.64 mmol/kg
Kaolinite . . . . . .1.40 mmol/kg
Oligoclase . . . . . 0.44 mmol/kg
Fluorite . . . . . 0.20 mmol/kg
Sphalerite . . . . . 0.11 mmol/kg
Illite/Sericite . . . . . 0.032 mmol/kg
Chalcopyrite . . . . . 0.029 mmol/kg
Goethite . . . . . -7.40 mmol/kg
Silica . . . . . -2.89 mmol/kg

Pyrite oxidation: the chemical model

$FeS_2 + 3.5O_2 + H_2O \rightarrow Fe^{2+} + 2SO_4^{2-} + 2H^+$

Pyrite + air + water ➜ acid ferrous sulfate soln

$FeS_2 + 3.75O_2 + 2.5H_2O \rightarrow Fe(OH)_2^+ + 2SO_4^{2-} + 3H^+$

Pyrite + air + water ➜ hydrolyzed ferric sulfate soln

$FeS_2 + 3.75O_2 + 3.5H_2O \rightarrow Fe(OH)_3\downarrow + 2SO_4^{2-} + 4H^+$

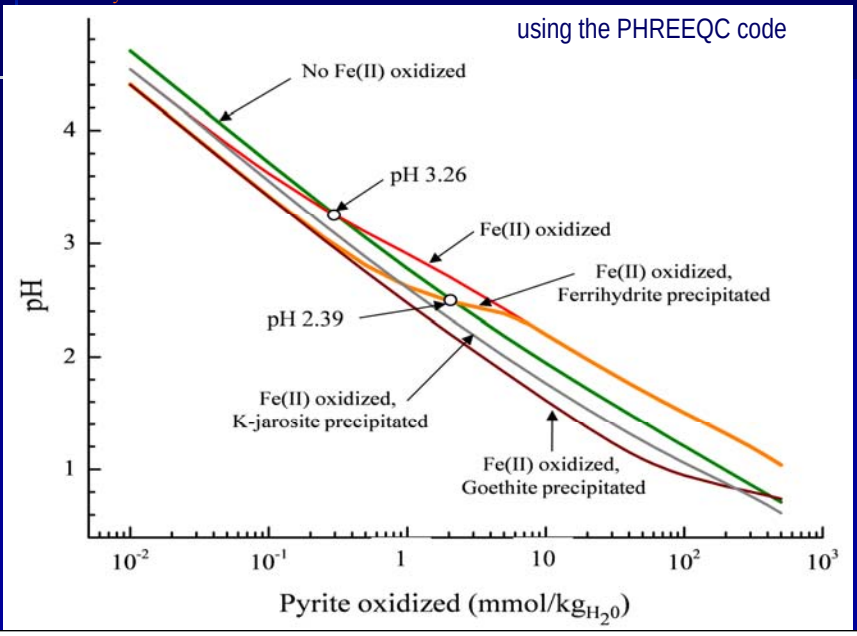
Pyrite + air + water ➜ sulfuric acid + iron ppt

Example 1.3 Pyrite oxidation: the graphical model (new insights)

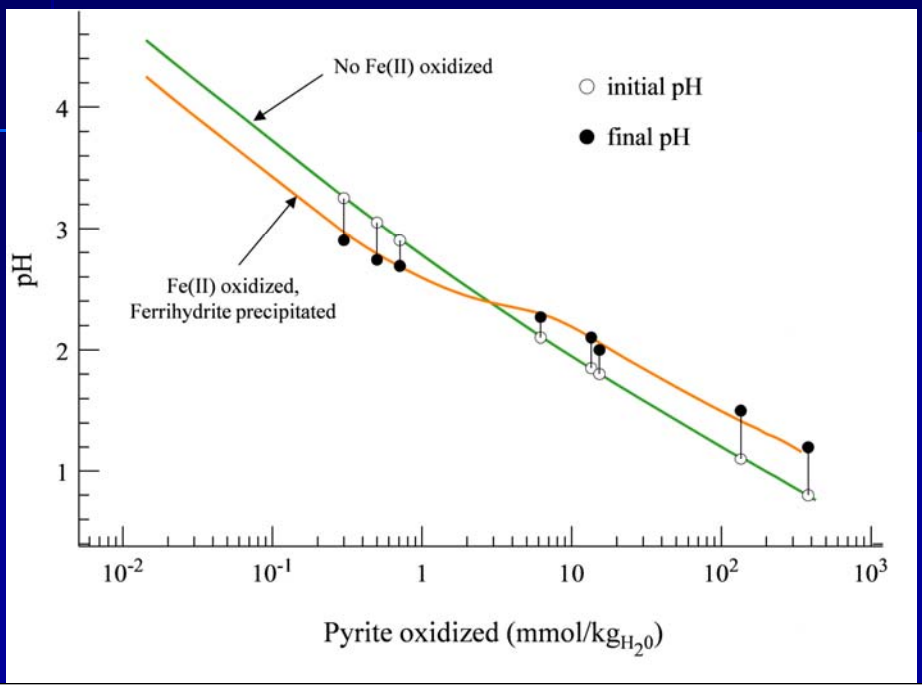
Simulation of pyrite+ O_2 + H_2O ⟶ Fe(II/III) + H_2SO_4

from Nordstrom and Campbell (2013) Modeling low-temperature geochemical processes, vol. 5, ch. 2, Treatise on Geochemistry

using the PHREEQC code



Example 1.4. Model testing with field data: confirmation

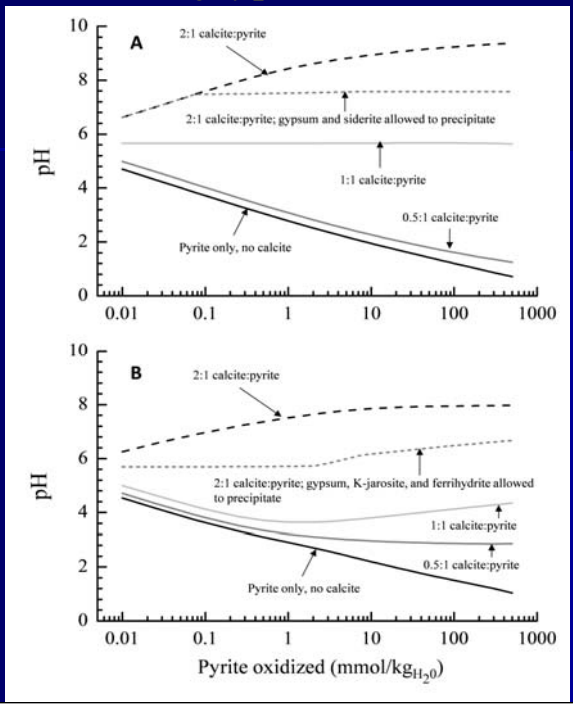


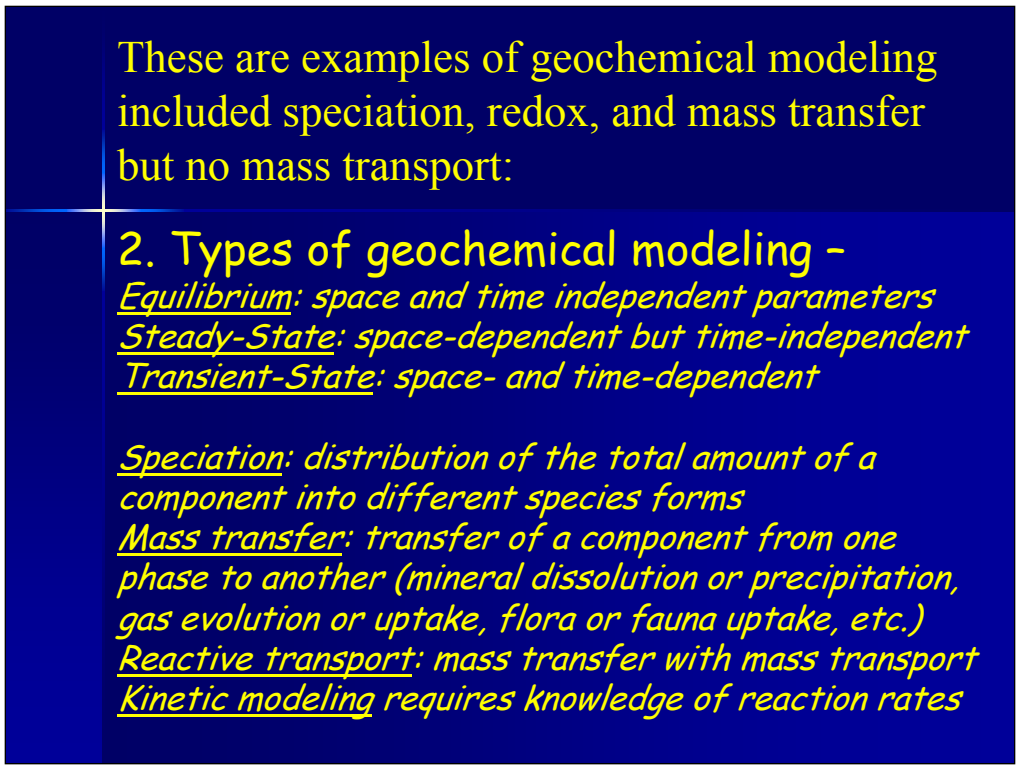
Example 1.5. Modeling of pyrite + calcite dissolution

Without oxidation of dissolved Fe

With oxidation of Fe

from Nordstrom and Campbell (2013) Modeling low-temperature geochemical processes, vol. 5, ch. 2, Treatise on Geochemistry (in press)



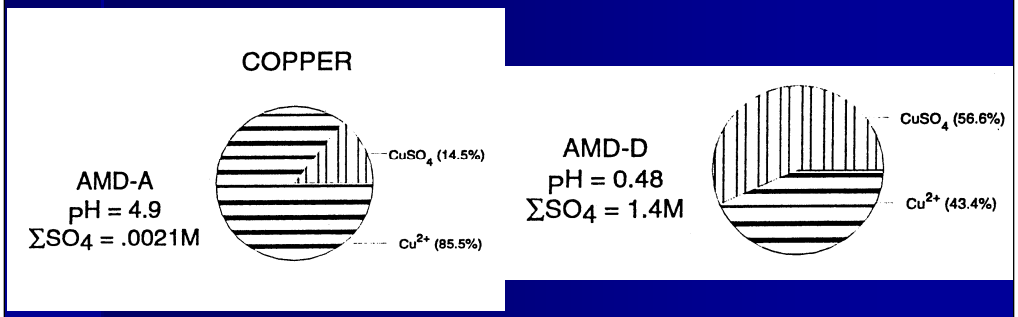


If kinetic modeling is included, everything gets very complicated and it becomes much easier to tweek code calculations so that the results come out anyway you want. Then the assumptions become very important and must be spelled out.

Example 2.1 Types of geochemical modeling: Aqueous speciation

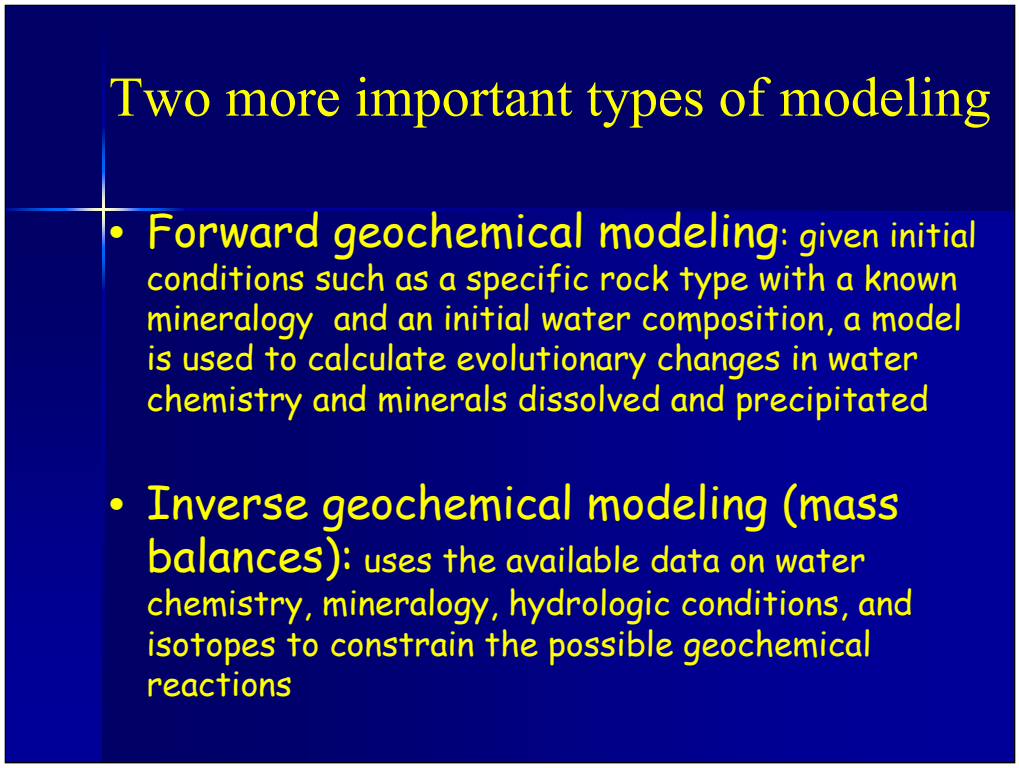
(ion-association model; allows calculation of SIs)

- AMD-A, Cu = 0.09 mg/L AMD-D, Cu = 290 mg/L



AMD-A: $Cu/SO_4 = 0.45$

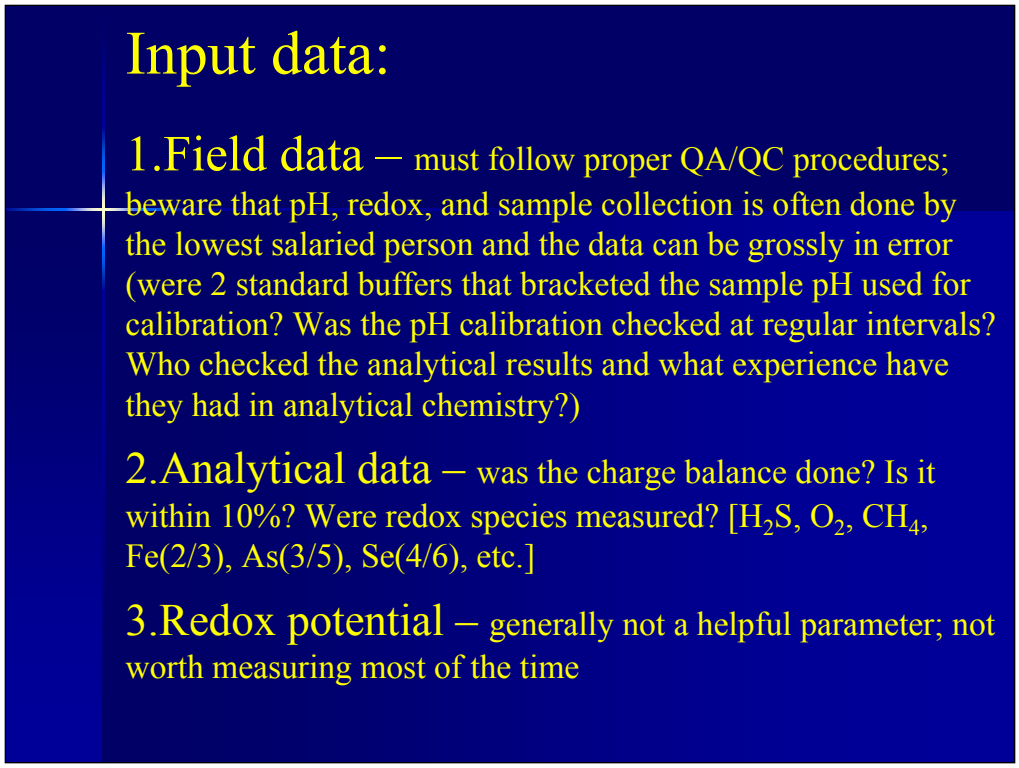
AMD-D: $Cu/SO_4 = 0.002$ much higher amount of Cu-SO_4 complexing



It is important to distinguish between these 2 types of modeling. Forward modeling might be used when there is little or no site data or lab data. Inverse modeling makes maximum use of field data. Forward modeling puts a huge burden on the modeler to know a great deal about geochemical, hydrological, and microbiological processes and site conditions. It is much harder to have confidence that forward modeling is useful by itself – far too much uncertainty. The optimal approach is to collect as much field data as possible and then fill in aspects of the geochemistry with forward modeling where there is insufficient data (see Glynn and Brown, 2012). IMHO, for regulatory purposes, if insufficient site data exists to do inverse modeling then any modeling is likely to be highly unreliable.

There are also different methods for calculating activity coefficients

The important point here is that if the ionic strength of the waters to be modeled is greater than 1 molal then the Pitzer ion-interaction model must be used. The Pitzer model does not have all the parameters needed for all metals yet but its database is improving.



Remember it is absolutely imperative to measure any important redox species for the purposes of chemical modeling. It is <u>not possible</u> to take a redox potential measurement with an electrode and determine the concentration of As(3/5) or U(4/6) – maybe Fe(2/3) under optimal conditions, but not recommended (especially because it is so easy to measure directly).

Why is direct measurement of aqueous redox species essential to geochemical modeling?

Element	Reduced form	Oxidized form
Fe	Fe(II), soluble	Fe(III), insoluble
As	As(III), soluble + more toxic	As(V), insoluble
Se	Se(IV), insoluble	Se(VI), soluble + more toxic
Cr	Cr(III), insoluble	Cr(VI), soluble + much more toxic

If Fe is a dominant cation in the sample, then the redox species must be analytically determined for charge balance as well as any speciation modeling.

[These solubility generalizations do not hold for all situations]

Saturation indices can be helpful guides as to whether a mineral could precipitate or not. Data from 1380 samples of AMD throughout the western US. Results show unreasonable supersaturation (up to 3.5 orders of magnitude). (Nordstrom, 2011)

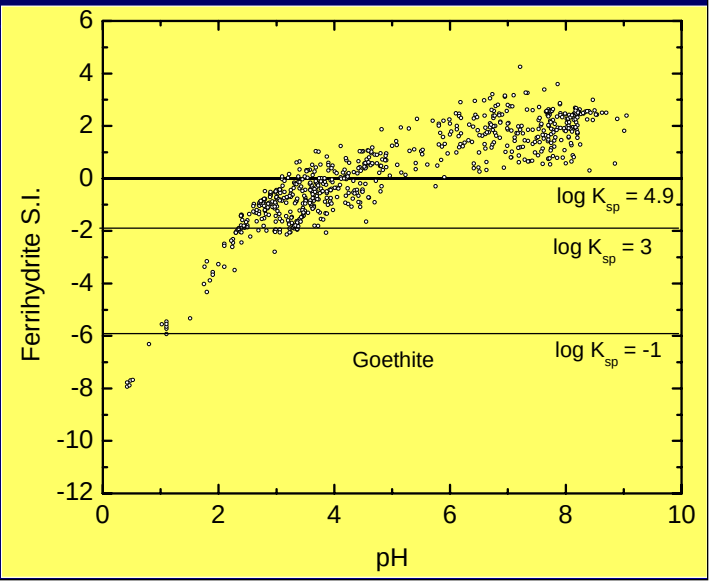
Ferrihydrite represents freshly precipitating $Fe(OH)_3$

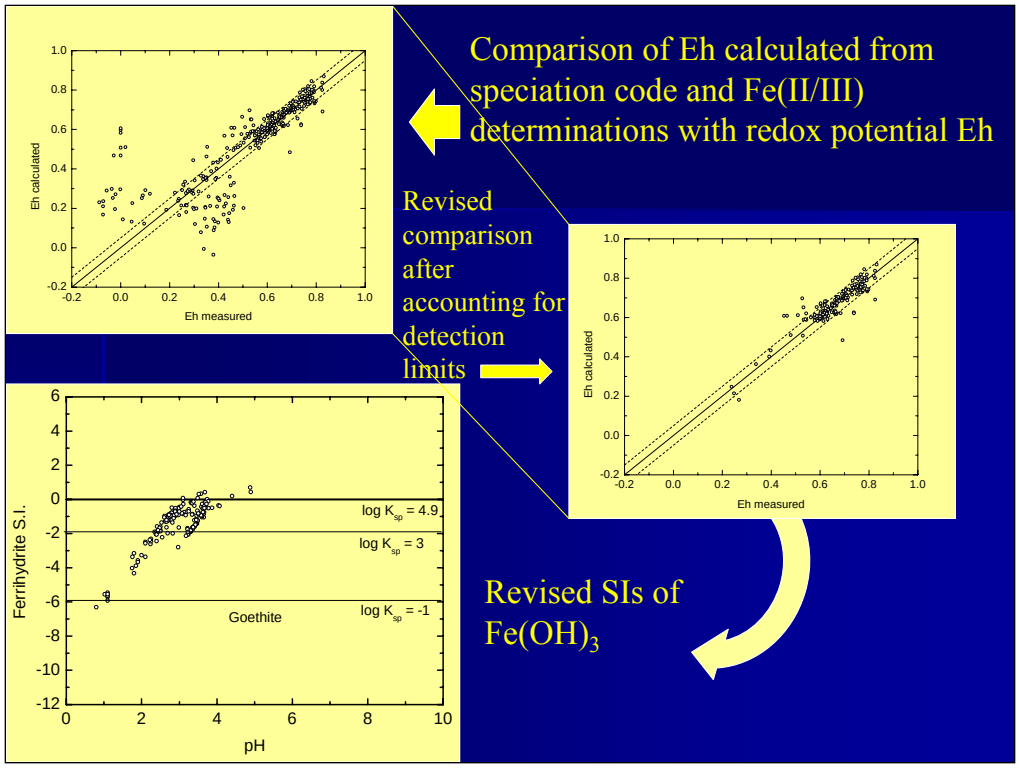
SI = saturation index

$$SI = \log\left(\frac{IAP}{Ksp}\right)$$

IAP = ion-activity product
Ksp = solubility-product constant

SI = 0, equil. soly.
SI > 0, supersaturation
SI < 0, undersaturation




Comparison of Eh calculated from speciation code and Fe(II/III) determinations with redox potential Eh
Eh calculated
Eh measured
Revised comparison after accounting for detection limits
Eh calculated
Eh measured
Ferrihydrite S.I.
pH
log K_sp = 4.9
log K_sp = 3
log K_sp = -1
Goethite
Revised SIs of Fe(OH)3

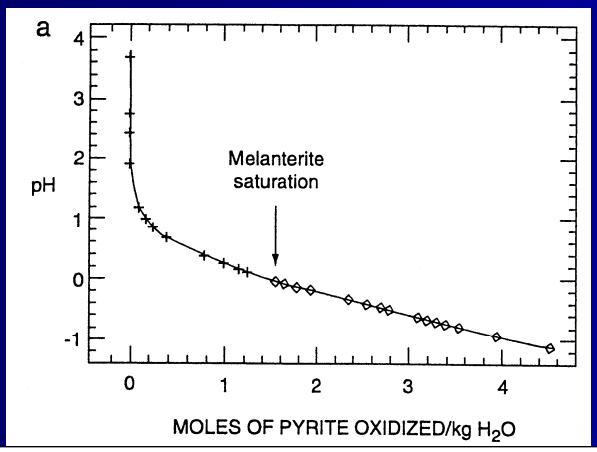
Chemical database – a) thermodynamic data and b) kinetic data

several to choose from; also should follow QA/QC procedures; has the database and the code been checked out against well-established independent data?? Can the modeler demonstrate that his calculations compare well with test cases or examples that have been done by other codes? Tests should include speciation, redox potential, saturation indices, mineral dissolution and precipitation rates if kinetics are involved, and reactive transport if relevant. What database was used and how does the modeler know that it is reliable?

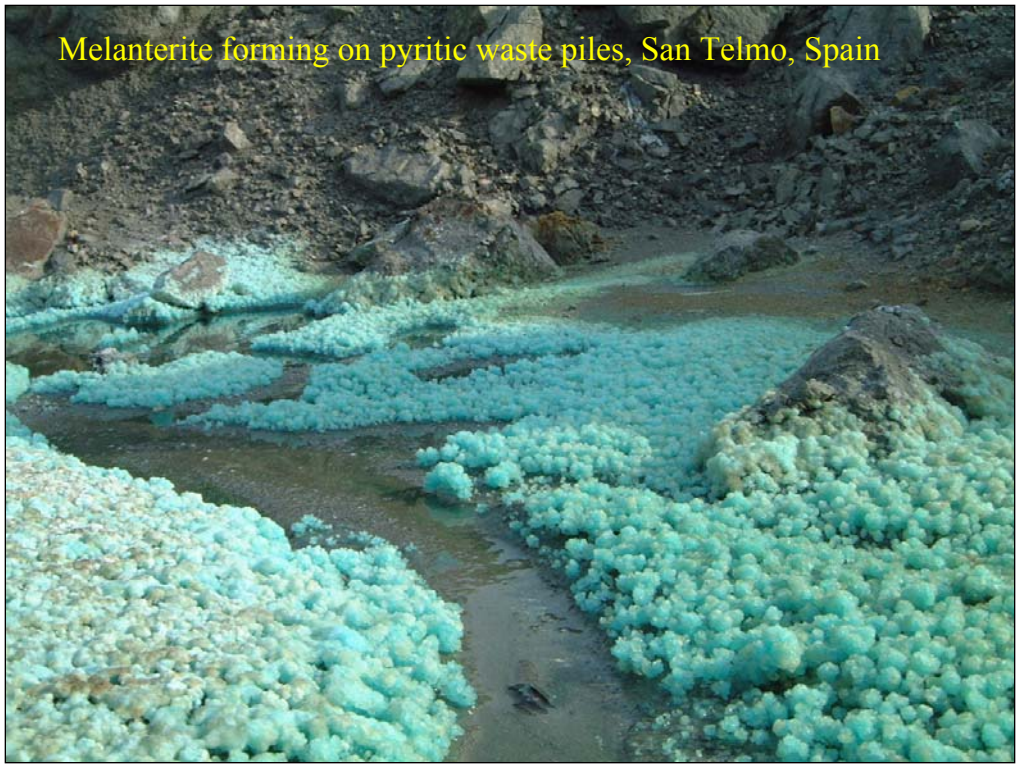
Note: no thermodynamic or kinetic database is necessary for mass balances

Example 2.2 Types of geochemical modeling: Mass transfer

Pyrite dissolution with melanterite precipitation (melanterite = $FeSO_4 \cdot 7H_2O$) using PHREEQPITZ code

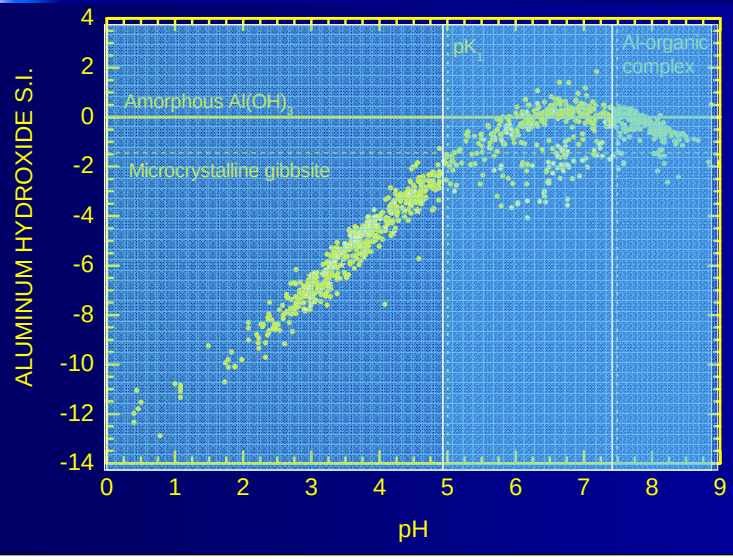


This computation tells us that melanterite should be forming from waters that have negative pH values. We have a confirmation of this from field data collected at Iron Mountain Mines Superfund site.


Melanterite forming on pyritic waste piles, San Telmo, Spain

A stalactite of melanterite was found underground at Iron Mountain Mines with water dripping from its tip with pH of -0.7 (the beaker is 2 liters for scale). Also this water had a temperature of 35°C underground and when it was brought outside to cooler temperatures of about 22°C, about a third of the water crystallized to melanterite, suggesting it was near equilibrium conditions.

Another example of mass transfer: Al precipitation, very common in mine drainage – from modeling we have been able to understand and generalize the geochemical behavior of Al in aqueous systems: precipitation occurs at pH ≥ 5.0 (pK_1=5.0)



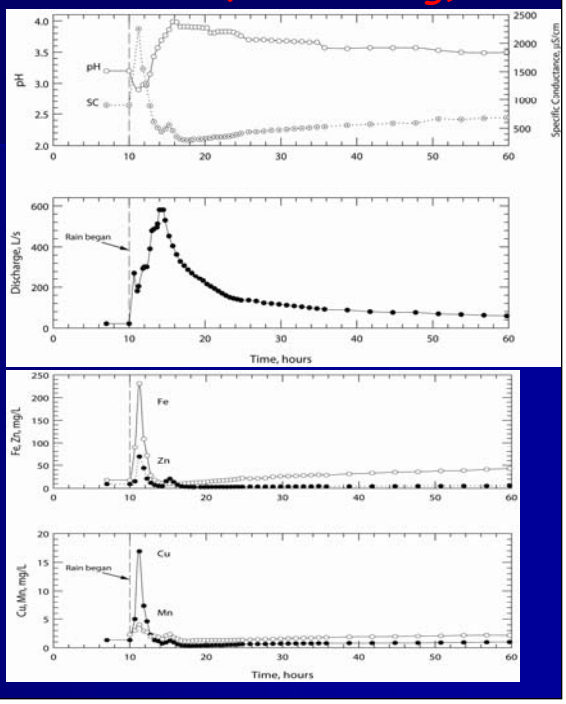

From Iron Mountain
From Paradise Portal, Silverton, CO
From Bog Iron Deposit at Opl

2.5 Example of : transient state *(no modeling)*

Contrary Ck, Virginia [MS thesis, T.V. Dagenhart, 1980]

Transient signal from flushout of soluble salts on tailings piles from a rainstorm

Can we predict this event? Qualitatively – in the sense that we know it happens, but not quantitatively

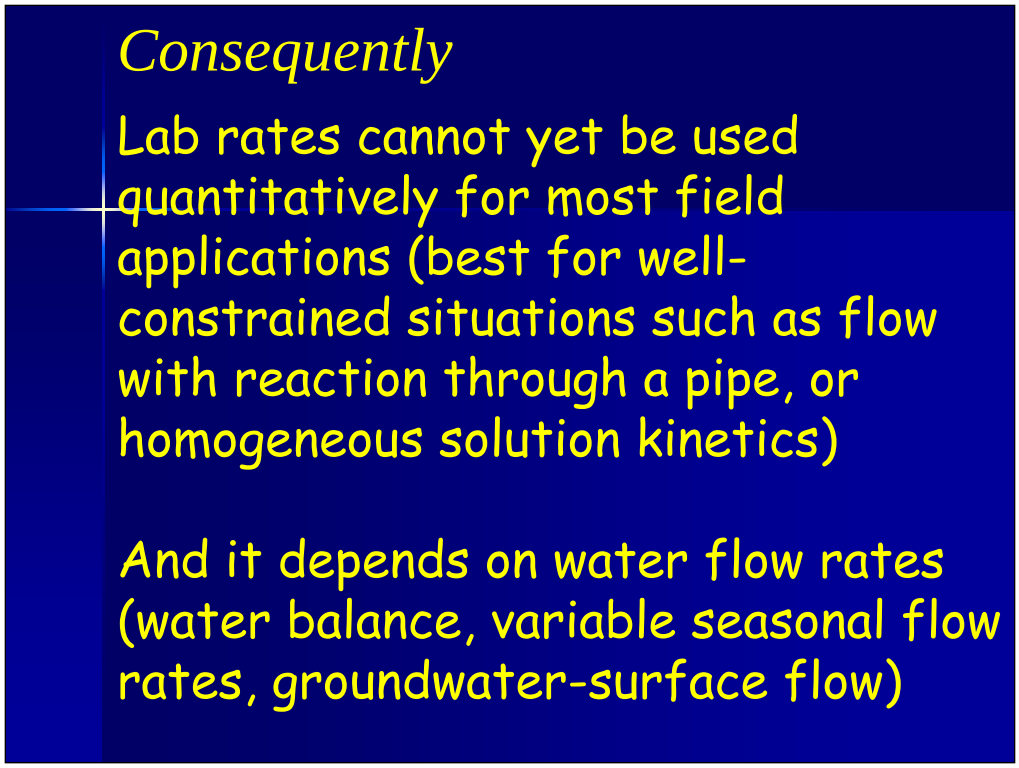


What about modeling rates??

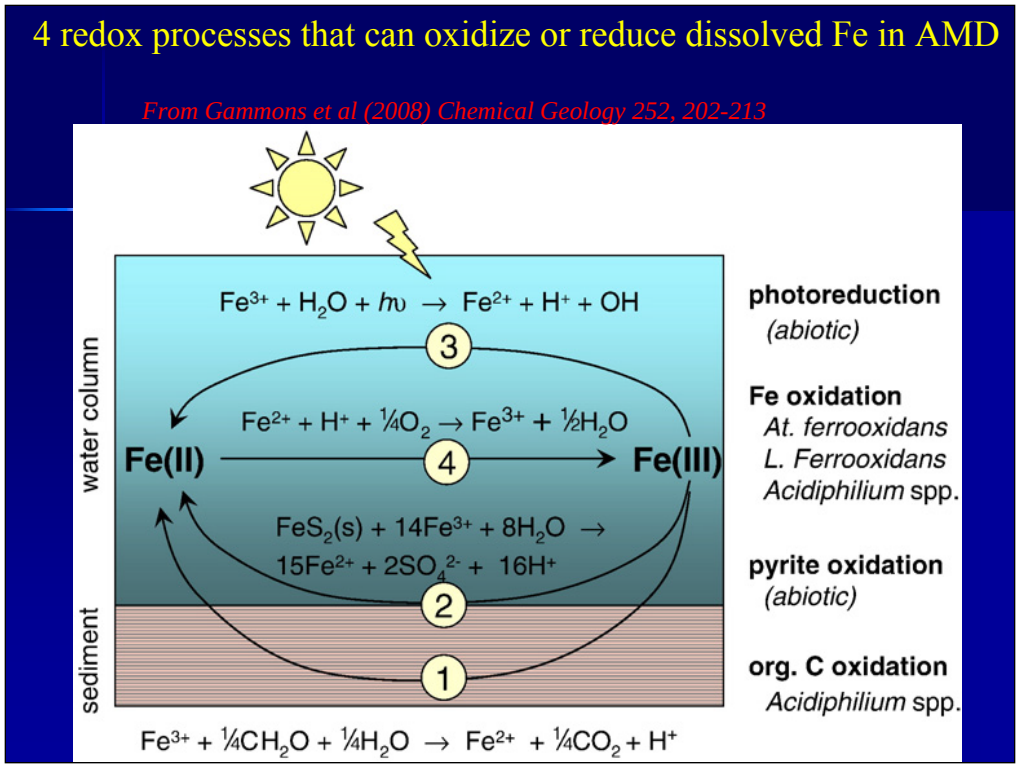
Comparison of lab-based mineral dissolution rates and field-based weathering studies of catchments have found that there is little agreement between the two. Usually there are orders of magnitude difference.

Numerous reasons have been given:

- Lab samples were ground and much more reactive
- Field samples have developed clay or silica coatings
- Reactive surface areas in contact with water are not known for field studies (surface areas and exact flow paths are unknown)
- Temperature and gas gradients occur in the field
- Organic matter and microbial activity affect weathering in the field in ways that are difficult to determine quantitatively
- Wet/dry cycles and seasonal changes occur in the field
- Residence time in weathering zone is much longer in the field and is not often measured
- It is not widely recognized that a lab-based study is a "generation" or "production" or "reaction" rate, whereas a field measurement is usually a "flux" or "transport" rate



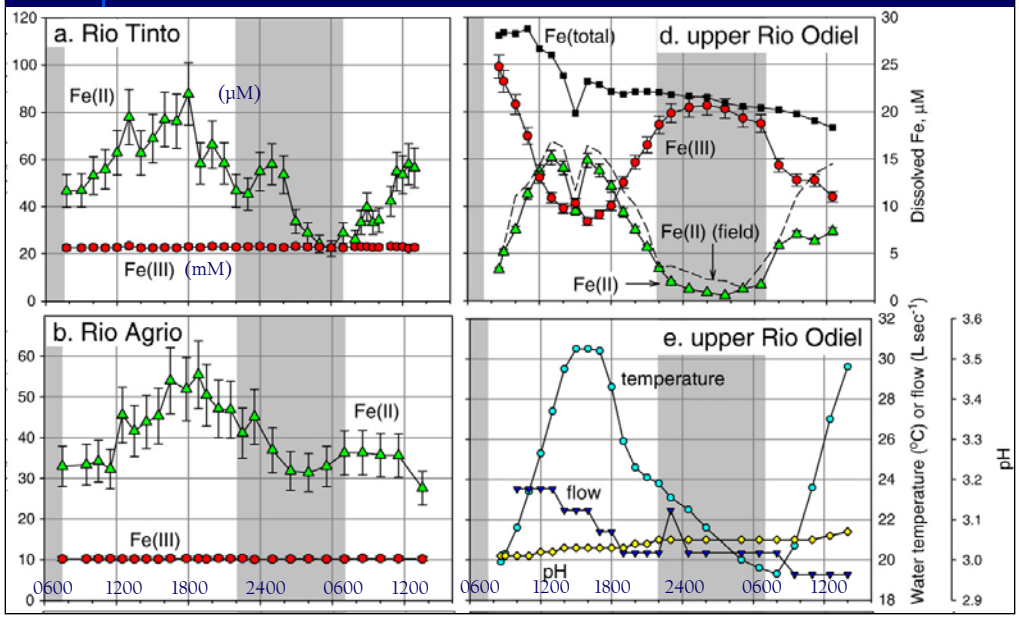
We are currently engaged with the Iron Mountain Mines Superfund site, determining the rates of iron oxidation and precipitation in a diversion pipeline and the geochemical and microbiological factors. Stay tuned.



Here are the 4 processes that can oxidize or reduce dissolved Fe in AMD. One can measure the net result (as Gammons et al did for the Rio Tinto, Rio Agrio, and Rio Odiel) and recognizing that microbial oxidation and photoreduction for surface waters are predominant in controlling Fe(II/III) concentrations and measuring concentrations over time during night and day, these investigators were able to determine each rate. This study is an excellent review and application of the state-of-the-art with respect to Fe(II) oxidation and Fe(III) reduction in AMD.

2.7 Examples of microbial oxidation and photoreduction in the Rio Tinto, Rio Agrio, and Rio Odiel, Spain

From Gammons et al (2008) Chemical Geology 252, 202-213



The fact that the Fe(II) concentrations never go to zero (except in the upper Rio Odiel) suggests there may be some Fe reduction in the sediments and diffusion of Fe(II) into the water column. Although lab and field rates are somewhat similar, there are some differences and if you want to know field rates, you have to measure them. One obvious source of uncertainty in any prediction is how much cloud cover will there be?

Can geochemical models be used to deterministically predict future scenarios at potential mine sites? Some summary comments….

- 50-100% of mine sites exceeded their predicted water quality conditions (see Kuipers et al., 2006)

"The computational capabilities of today's codes and advanced computers exceeds the ability of hydrogeologists and geochemists to represent the physical and chemical properties of the site or to test the outcome of the model." Maest et al. (2009)

Can reactive transport models be used to deterministically predict future scenarios at potential mine sites? Some summary comments….

"Tempting as it will be to government bureaucrats to continue the use of models, the predictive models for the long-term quality of water in abandoned open-pit mines should themselves be abandoned." Pilkey and Pilkey-Jarvis (2007)

"Just as in other modeling arenas we have discussed, accurate prediction of future water quality is a fantasy supported by a hyperreligious faith in the predictive power of numbers." Pilkey and Pilkey-Jarvis (2007)

Can reactive transport models be used to predict deterministically future scenarios at potential mine sites?

"Reactive transport models cannot solve the problem of the apparent discrepancy between laboratory and field rates by themselves…." Steefel et al. (2005)

"… the reactive transport modeling can be used to narrow down the possible explanations for the overall rates observed in the field." Steefel et al. (2005)

"Another possible approach is to choose field sites where the transport rates can be modeled accurately and deterministically because gross physical heterogeneities are absent." Steefel et al. (2005)

Do such places exist? Not at hardrock mine sites!

Validation & Verification

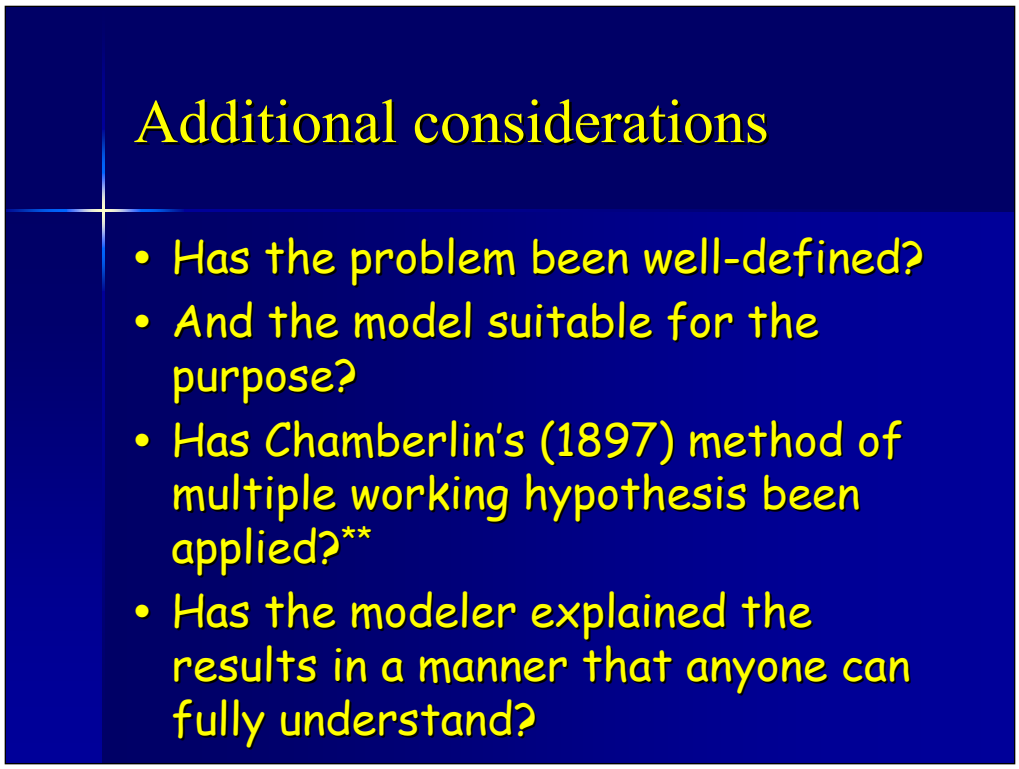
- "Does good agreement between a model result or prediction and observational measurements mean the model is correct?
- No, for 3 possible reasons, (1) if model parameters are not independent from the measurements they are being compared to, they should agree regardless of the correctness of the model, (2) if the measurements are in error then both the measurements and the model could be in error, and (3) the model results might agree with reliable measurements for the wrong reasons.
- Does poor agreement between a model result and observations mean the model is incorrect?
- No, for similar possible reasons, (1) if the measurements are unreliable, the model may still be correct, (2) model calculations could be in error whereas the conceptual model could be correct, and (3) the criteria for what constitutes good and poor agreement may be incompatible with the limitations and uncertainties of the model." [Nordstrom, 2012] The criteria for agreement could be made too broad or too confined. Hence, a model could be validated or invalidated according to preconceived agendas.

Recommendations

- Don't use the word validation with respect to a scientific model; it doesn't apply [see Nordstrom, 2012]
- If someone says a model has been validated, ask him/her to invalidate it (it can always be done); then have them draw their own conclusions
- Remember: models are not unique or exact!

Recommendations

- It is the quality of the conceptual model that determines the usefulness and relevance of any modeling; the conceptual model needs peer review
- There will always be unknown factors that affect our confidence in modeling
- Computer codes in the regulatory realm must be transparent!
- Is it necessary to predict far into the future? Or is it better to use best available technology and protect the public and the environment through other means (liability)



**An appropriate corollary of Chamberlin's method is that if someone has presented what you feel is a best-case scenario prediction, have them do a worst-case scenario. If someone has "validated" their model, have them also invalidate it!

Conclusions?

With models we can constrain the possible explanations for our observations.

1. We cannot model without observations.
2. The more observations we have the better will be our modeling.
3. With enough observations, we don't need a model.

"More broadly, it means recognizing that the amount of confidence someone expresses in a prediction is not a good indication of its accuracy – to the contrary, these qualities are often inversely correlated."
(Silver, 2012)

"It isn't what we don't know that causes the trouble, it's what we think we know that just ain't so." Will Rogers