A probabilistic analysis of meteorically altered δ13C chemostratigraphy from late Paleozoic ice age carbonate platforms
Dyer et al.

1. SUPPLEMENTARY INFORMATION

1.1. Bayesian Inference

For the simplest models, this Bayesian inference problem has a complex analytical solution, but for most models an analytical solution is intractable. Additionally, in these more complex models, as the dimensions (parameters) of interest in the model space grow, brute force exploration of the entire space grows exponentially with each dimension. The Markov Chain Monte Carlo (MCMC) algorithm is a probabilistic numerical approach for solving this high dimensional problem for complex models (Metropolis et al., 1953; Hastings, 1970). The algorithm is designed to hone in on and randomly walk through regions of high probability in the parameter space to describe the posterior distributions for each parameter.

The general numerical approach is to draw a set of model parameters from their prior distributions based on the current state of the model. The relative likelihood that the observations correspond to current or proposed model state is used to accept or reject the proposal. Accepted parameters are stored, and rejected parameters return the model to the previous state. This process is then repeated many times. The MCMC sampler used for Bayesian inference in this work is a Metropolis-Hastings algorithm that sets the probability for accepting a proposal to the ratio of the likelihood of the proposal to the likelihood of the current state (Metropolis et al., 1953; Hastings, 1970; Patil et al., 2010). Therefore, parameter proposals that fit the data better are accepted, and proposals that are a worse fit to the data are still sometimes accepted based on their relative likelihood to the current state. To improve the speed of convergence on the posterior distribution, after 200 accepted proposals the sampler scales (tunes) the proposal distributions based on the rate of accepted proposals. High acceptance rates lead to expansion of the parameter proposal distributions to enable sampling of a wider region of parameter space, and low acceptance rates leads to a more focused proposal distribution. In theory, infinite samples of accepted proposals fully solve the analytical Bayesian inference problem. However, convergence to a posterior estimate often requires many fewer samples, and can be evaluated by inspecting the trace of accepted parameters. We sampled the posterior distributions for 100,000 accepted parameter proposals for each inference performed in this study. The traces (series of accepted values) were inspected for signs of 'good mixing' to indicate that the algorithm is properly sampling the full region of high probability. Well mixed traces will have decreasing amounts of autocorrelation for higher lag intervals, and a running mean should converge to a single value.

1.2. Testing the Bayesian Process model on Model Generated Data

Test datasets with known parameters can be generated from the forward model and run through the Bayesian inference model to determine the success of the inference model at recreating the true parameter values. Figure 2 illustrates the results of the MCMC algorithm on three datasets generated from the model and normally distributed random noise. Test A and B are run with diagenetic durations that fall within the prior distribution for this parameter, which represents some geologic limits on exposure length. The posterior distributions for each test are nearly identical to the prior proposal, which indicates that the dataset does not inform that prior information very strongly. However, the inference on the data provides a range of reaction rates that are as good as the prior distribution used for the duration of diagenesis. For a third test (C), the model was run for a longer duration that falls beyond the prior distribution for duration of diagenesis. In this case, the MCMC algorithm fails to capture the true value within the posterior distribution for both reaction rate and duration of diagenesis. However, the posterior distribution for the ratio of advection to reaction in all three tests includes the true parameter used to generate the model. Therefore, the length scale and curvature of the reaction front do an acceptable job at informing the inference model about this advection to reaction ratio, regardless of a bad prior for the duration of diagenesis. Since the reaction rate serves to convert this ratio into a geologically meaningful carbonate mass flux, these estimates will only be as good as the posterior distribution on the duration of diagenesis. In the failed case of test C, the true duration of diagenesis is roughly four times longer than the posterior distribution, and thus reaction rates are also overestimated by four times.

1.3. The Variation in Advection to Reaction

The dissolution of carbonate in the vadose zone primarily is driven by the mixing of fluids that have different CO_2 concentrations (Wigley and Plummer, 1976; Drever, 1982; Gulley et al., 2014). CO_2 from the vadose zone and soil accumulates and drives the expansion of cave systems in carbonates. When caves are well ventilated to the surface,

this effect is minimized. However, the air in cave systems with large seasonal variations in outside air temperature (sub-tropical) tend to stagnate during summer months due to density contrasts in the air masses, and CO_2 from the fluids moving through the vadose and soil zones can build up in the cave system (Gulley et al., 2014). Additionally, organic acids and supercharged CO_2 fluids in the soil zone can lead to rapid dissolution of the upper carbonate surface from above, resulting in the common pitted karst terrain topography (Ford and Williams, 2013). The close link between CO_2 and dissolution suggests that the high weathering fluxes of carbonate in the Leadville region resulted from some combination of seasonality, a large local terrestrial biosphere, and relatively fast vertical fluid velocities through the phreatic zone. Further interpretation will require new isotopic observations from the carbonates in modern karst terrains where these environmental conditions are well known.

1.4. Duration of Sedimentary Hiatus

In addition to recording information about the mass and isotopic fluxes through the carbonate platform, the isotopic excursion also depends on the duration of diagenesis and rate at which reactions between the fluid and rock occur. If the reaction rate between carbonates and meteoric fluids determined here are realistic, then the geologic observations of isotopic excursions in other stratigraphic settings can be used to estimate the duration of time missing at exposure surfaces. The carbonate geologic record is filled with instances of parasequence bounded meter-scale meteoric diagenesis at exposure surfaces (Elrick and Scott, 2010; Dyer and Maloof, 2015), and placing limits on the duration of exposure can test hypotheses linking the carbonate parasequence architecture to orbital forcings on climate (Goodwin and Anderson, 1985; Grotzinger, 1986; Read, 1998; Goldhammer et al., 1990). The rates determined above are a first step towards establishing a tool for estimating the missing time at any meteorically altered exposure surface in carbonate rocks.

2. Supplementary References

Drever, J., 1982, The Geochemistry of Natural Waters: .

Dyer, B. and Maloof, A. C., 2015, Physical and chemical stratigraphy suggest small or absent glacioeustatic variation during formation of the Paradox Basin cyclothems: Earth and Planetary Science Letters, v. 419, p. 63–70.

Elrick, M. and Scott, L., 2010, Carbon and oxygen isotope evidence for high-frequency (104-105 yr) and My-scale glacio-eustasy in Middle Pennsylvanian cyclic carbonates (Gray Mesa Formation), central New Mexico: Palaeogeography, Palaeoclimatology, Palaeoecology, v. 285, p. 307–320.

Ford, D. and Williams, P. D., 2013, Karst hydrogeology and geomorphology: John Wiley & Sons.

Goldhammer, R., Dunn, P., and Hardie, L., 1990, Depositional cycles, composite sea-level changes, cycle stacking patterns, and the hierarchy of stratigraphic forcing: examples from Alpine Triassic platform carbonates: Geological Society of America Bulletin, v. 102, p. 535.

Goodwin, P. and Anderson, E., 1985, Punctuated aggradational cycles: a general hypothesis of episodic stratigraphic accumulation: The Journal of Geology, p. 515–533.

Grotzinger, J., 1986, Cyclicity and paleoenvironmental dynamics, Rocknest platform, northwest Canada: Geological Society of America Bulletin, v. 97, p. 1208–1231.

Gulley, J., Martin, J., and Moore, P., 2014, Vadose CO2 gas drives dissolution at water tables in eogenetic karst aquifers more than mixing dissolution: Earth Surface Processes and Landforms, v. 39, p. 1833–1846, doi:10.1002/esp.3571.

Hastings, W. K., 1970, Monte Carlo sampling methods using Markov chains and their applications: Biometrika, v. 57, p. 97–109.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., 1953, Equation of state calculations by fast computing machines: The journal of chemical physics, v. 21, p. 1087–1092.

Patil, A., Huard, D., and Fonnesbeck, C. J., 2010, PyMC: Bayesian stochastic modelling in Python: Journal of statistical software, v. 35, p. 1.

Read, J. F., 1998, Phanerozoic carbonate ramps from greenhouse, transitional and ice-house worlds: clues from field and modelling studies: Geological Society, London, Special Publications, v. 149, p. 107-135.

Wigley, T. and Plummer, L., 1976, Mixing of carbonate waters: Geochimica et Cosmochimica Acta, v. 40, p. 989–995.

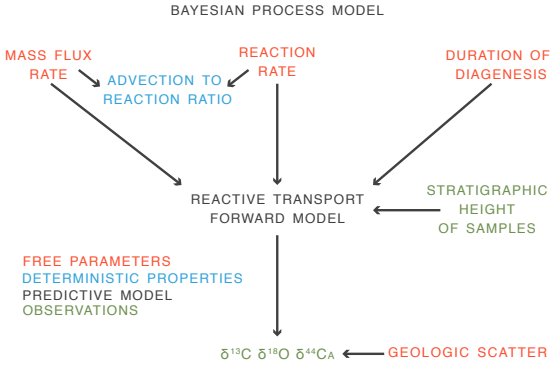

Figure DR1: This figure depicts a process diagram for the Bayesian inference problem. The forward reactive transport model is used to predict isotopic excursions, which are then compared to the observations from the middle Carboniferous karst terrain. The Markov Chain Monte Carlo sampling algorithm proposes new parameter values for each of the free parameters, and accepts or rejects the parameter set based on the likelihood that the observed isotopic data was generated by the model with those specific parameters.

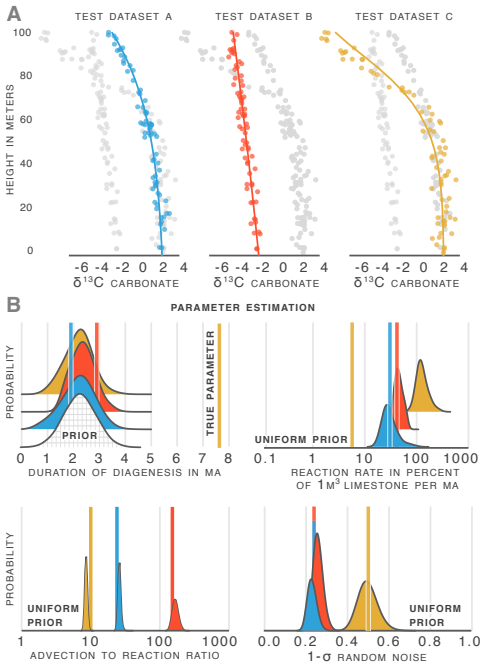

Figure DR2: Datasets generated by the forward model with known parameters are passed to through the Bayesian inference model to explore the parameter estimation and its pitfalls. A. Carbon isotopic profiles from three test datasets generated by the forward model with normally distributed random noise added to each datapoint. B. Prior (box swatch) and posterior (filled color) distributions for the duration of diagenesis, reaction rates, advection to reaction ratios, and random noise. The colored vertical lines highlight the true value of each parameter used to generate the test dataset.

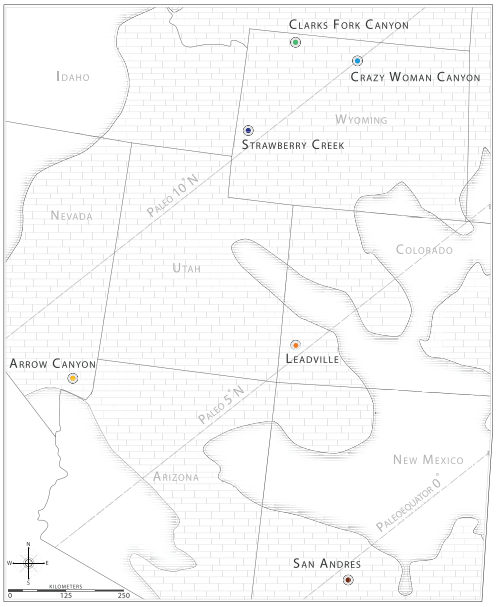

Figure DR3: Regional map of the western U.S. during the Mississippian subperiod. Blank regions represent highlands or regions with no Mississippian sediment preservation, and block texture represents marine carbonate deposition. Diagonal lines are paleolatitude (**?**).