Reply letter, reviewer #2

Thanks to the authors for clarifying their manuscript.

I have some additional minor points for clarification, that I unfortunately did not notice during the first review round. It should help the reader who want to reproduce your work or apply your method.

Figure 1: A, B, C, D called in text and caption but missing from the figure, though obvious.

corrected

Line 147: though it does not affect the methodology presented in the paper, wouldn't it make more sense to define i(x) as the indicator of m(x)>t rather than z(x)>t ? It would also further support inferring m(x), unless the prior of m(x) ad r(x) is better characterised than the prior of z(x).

Indeed both options can be made

- i(x) as indicator of m(x)
- i(x) as indicator of z(x)

In reality the r(x) is more complex, as noted in the text and consisting of "confounding elements" related to any modification of the ore body, so in a real context r(x) will not be a simple additive noise term.

Change, line 117: "…and hence the noise term in this simple example is used to develop a methodology"

Line 159: based on a minimum estimated volume of ore?

Changed: "The question we will address is: what is the optimal sequence of data acquisition that best informs "mine" vs "do not mine" decision, based on a mineable volume exceeding some minimum threshold?"

Line 226: I am a bit confused by b' – do you mean b(s_{t+1}) proportional to L(o_{t+1} given s_{t+1} and a_t) times b(s_t) ? b' is only used later in the pseudo-algorithm

This a common notation used in AI

The b' notation here is the posterior so it includes the condition | o_{t+1}

Change line 227: "Note that $b'(s_{t+1})$ is AI notation for a posterior $p(s|o)$, where $p(s)$ is the prior"

Figure 4: use t and t+1 rather than t-1 and t to be consistent with the caption and manuscript description?

Corrected

Line 333: it looks like r(x) is missing – do you mean f(m,r , o)=f(o)*f(m given o)*f(r given m) ?

Lines 335 to 337: I am confused by the notation versus the description of f(m given o) (conditional belief in my understanding) and f(m,o) (joint belief in my understanding), can you check this?

Line 342: posterior f(m given o) ?

Line 349 and 351: I am not sure to understand how the decomposition of o_t as o_tm+o_tr and the determination of o_tm by m(x) lead to these weight equations.

We apologize for this issue; it seems something happened when converting to Word from google doc.

Here is the section with correct equations

$$o_{1:t} = \{o(x_\alpha), x_\alpha = 1, \dots t\}$$

The observed measurements are dependent upon both random functions, $m(x)$ and $r(x)$, hence a traditional conditional simulation cannot be directly applied. Instead, we formulate this problem as a hierarchical Bayes' problem by factoring the joint distribution into

$$f(m(x), r(x)|o_{1:t}) = f(m(x)|o_{1:t}) \times f(r(x)|m(x), o_{1:t})$$

Samples are generated from this distribution hierarchically by first drawing a sample from the distribution over $m(x)$ and then using the resulting sample to draw from the conditional distribution over $r(x)$. We model the belief $f(m(x)|o_{1:t})$ as a particle set and update it using an importance resampling particle filter (Del Moral 1996, Liu et al. 1997). The conditional belief $f(r(x)|m(x), o_{1:t})$ is modeled as a conditional Gaussian process.

A particle set is an ensemble of realizations of the state variable with a sample distribution approximating the true state distribution. The initial particle set is generated by first sampling an ensemble from the uniform prior distribution. For an n particle set, this corresponds to an ensemble of $\left(m^i(x), r^i(x)\right), i = 1, \dots n$ where each particle is equiprobable.

When new information o_t is observed, the particle filter updates the belief by updating the ensemble such that the new particles are sampled according to the posterior distribution $f(m(x)|o_{1:t})$. To do this, a posterior weight is calculated for each particle according to Bayes' rule as

$$w^i \propto f(o_t|m(x), o_{1:t-1})$$

Note that each particle is treated as equiprobable in the particle set, so the prior probability is dropped in the above expression. The observed measurement o_t is determined by the sum of $m(x)$ and $r(x)$ at the location of the measurement. We denote these values as o_t^m and o_t^r, respectively, such that $o_t = o_t^m + o_t^r$. Using this notation, we can decompose the particle weight function into

$$w^i \propto f(o_m^t|m(x)) \times f(o_r^t|m(x), o_{1:t-1})$$

Because the value of o_m^t is completely determined by $m(x)$, we can simplify this further to

$$w^i \propto f(o_t - o_m^t|o_{1:t-1} - m(x))$$

Line 370: insert 'At' before 'each time step t'

corrected

Line 422: do you mean n (size of the particle set) = 1E4 simulations? It becomes clear only in the conclusions that you refer to the total number of trial trajectories m. What is n, the size of the particle set?

What is meant is "trial simulations" or trial trajectories related to evaluation of the Monte Carlo Tree search,

Change, line 422: "We ran POMCPOW for 10,000 trial simulations (trajectories) per-step"

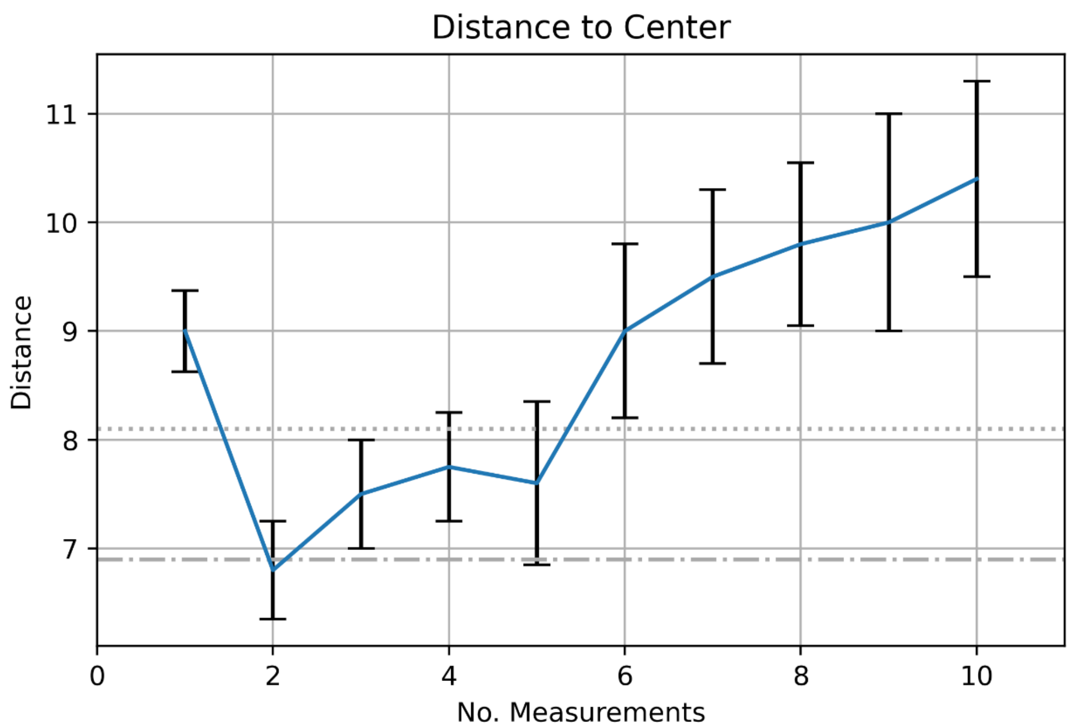
Did you define a maximum number of time steps 1) to restrain the depth of the search tree and 2) in the iterative process?

We define 25 measurements as a maximum but allow POMCPOW to search over full depth (selectively deepening search process)

Change, line 457: “We limited the agent to a maximum of 25 measurements”

Figure 16: would it be worth to plot the approximate radius of the orebody as an horizontal indicative dashed line?

Very nice suggestion, we changed all figures like these, here is an example



Here the dotted line is the maximum possible orebody. The fact that we step out further is because of the imperfect measurements we take on it