



Research and strategy for the land community.

To: Northeastern Minnesotans for Wilderness
From: Sonia Wang, Spencer Phillips
Date: 2/27/2018
Subject: Full results from the review of comments on the proposed withdrawal of lands from mineral leasing in the Boundary Waters.

Summary

Northeastern Minnesotans for Wilderness has asked Key-Log Economics to review scoping submitted to the Superior National Forest regarding withdrawal of certain lands from the federal mineral leasing program. In this initial review, we have identified comments as either favoring or opposing the withdrawal, and we have identified comments that reference economic concerns. Economic concerns included employment (jobs), income, tax revenue, school funding, future investment opportunity, and other factors.

Key-Log Economics, along with OB Consulting, used both a combination of human review and machine-learning tools to complete a review of all individual comments submitted to the Forest Service. Based on our review, we found that the public overwhelmingly supports the withdrawal.

In total, there were 81,032 comments submitted to the USFS directly. Of this total:

- 1,464 or 1.8% of the total comments oppose the withdrawal.
 - 1,219 oppose the withdrawal and made an economic argument.
 - 245 oppose the withdrawal for other reasons.
- 79,568 or 98.2% of the total comments support the withdrawal.
 - 48,965 support the withdrawal and made an economic argument.
 - 30,603 support the withdrawal for other reasons.

In addition to the individual comments, the Forest Service also received several petitions and many postcards. We reviewed those manually and report the results in the final section of this memo.

Methods

Submissions to the Forest Service could either be a unique comment, a form letter, petitions, or postcards. In our process we scanned every comment available from the USDA cloudvault sites and recorded the sentiment (oppose or support the withdrawal) for each submission and further categorized the comments read by the machine according to whether or not it mentions economic issues. For the postcards and petitions, we only recorded whether they support or opposed the withdrawal.

In the oppose/support classifications, some “unique” comments were actually attachments (a law case, an environmental study, etc.) to a parent comment. Those attachments received the same oppose/support classifications as the parent comment.

Table 1. Counts of Comments Received by Type and Mode of Analysis

	Records ^a	Comments Received ^b (i.e., persons' opinions represented)
Individual Comments (Unique letters & submissions of form letters) Reviewed via machine learning	81,032	81,032
Postcards Reviewed manually	143	4,056
Petitions Reviewed manually	1 (contained 5 petitions)	39,972
Total	81,176	125,060

Notes:

- a. A "Record" may be comprised of a single comment, or it can be a file with multiple comments in the form of bundled signatures on a form letter or signatures on a petition.
- b. "Comments Received" counts each single or bundled comment, and each signature on a petition, as an expression of an individual's opinion on the withdrawal

Machine Learning

Pre-Screening Comments for Form Letters

To identify form letters, we compared the first 100 characters of every comment. Based on this comparison, we separated the comments into 37 bins, each corresponding to one of 37 different form letters. We then manually classified form letter (and therefore the comments in its bin) as either for or against the withdrawal and either including or not including an economic argument. The 37 bins contained 76,715 comments in total. The remaining 4,317 comments were then classified using machine learning techniques as described below.

Pre-Processing & Annotation

Annotations are created using a custom data science platform where the classification model presents the annotator (OB Consulting) with documents it is least sure about. With each annotation, the model updates, allowing for much more efficient targeting. While discussion of economics requires a relatively small amount of annotation for accurate input, accurately determining for/against required significantly more annotation work due to similarity of words and logic (e.g. "I oppose the withdrawal of mining leases in the boundary waters" vs "I oppose mining leases in the boundary waters").

OB consulting manually annotated 1,732 comments. By manually classifying these comments, we can use these accurately classified comments to train the machine. Training the machine learning model is an iterative process where a model refines its predictions multiple times, until settling on the highest accuracy achievable with the training set.

Algorithm Results: Testing

The 1,732 comments in the annotated training set included comments from all classifications. From the full annotated training set, we separated 20% of the set out and used it as the testing set. These 346 comments were used to test the accuracy of the other 80% of comments, which were then used to train the machine model.

For the oppose/support classification, the model finished training with an accuracy of 94%, and an F-1 Score of 91%. An F-1 score is used to consider tradeoffs between precision and recall, and is likely a better metric of overall accuracy. These output metrics are calculated by looking at the binary results; in reality, the prediction scores are returned as a percentage where the model is more certain of its outputs for scores closer to 1 (positive classification) and 0 (negative classification). Scores near 50% are less certain, and scores hovering around 50% may be false positives/negatives.

The economics classification accuracy score returned 90% while the F1 score return 87%. Due to the variety of topics included in the classification (schools, jobs, economy, national security), the machine had a harder time identifying specifically what it needed to search for, in order to provide a True/False classification. Had single topics (such as education, jobs, national security) been classified individually, higher accuracy would be expected. However, the vast majority of comments against the withdrawal included some economics-oriented topic. This was noted during annotation and in the model outputs.

Other Considerations

There were around 1,200 comments that were not text-based PDFs, meaning they needed to translate image-based pdfs to text via Optical Character Recognition (OCR). While this process generally works and salvages most text, it is not perfect. While it is likely that it did not throw the classifications off to an extreme, it is likely that the 1,200 comments classified via the OCR method are not as accurate as the process above, but only by a small margin.

Out of the 81,032 total submissions reviewed by the machine, 130 were unprocessable due to the lack of text (e.g. images). These files generally had “(attachment)” designations in their filenames and have been excluded from the results.

Postcards and Petitions Analysis

There were 171 electronic files to analyze. These included the following:

- 1 file containing cover letters for 5 petitions, with each cover letter indicating the number of signatures for the petition and the sentiment expressed (all petitions, as it happens, were in favor of the lease withdrawal);
- 19 files containing only signatures for the petition, but often no indication of the organization to which the petition and the signatures belong;
- 143¹ files comprised of one or more postcards (with a total of 4,056 comments); and
- 9 files containing duplicates of petition cover letters, envelope scans, and other material not relevant for analysis.

¹ File 20171226153731.pdf contained multiple pages of petition signatures as well as one postcard. Because we deleted files that contained petition signatures, we duplicated this record so that the postcard was accounted for in the analysis.

For this analysis, all sentiments expressed were classified as a signature on a petition, or an individual comment, whether that comment came as a postcard or in another format.

For the petitions, we had used information in the cover letters to gauge the sentiment and the number of people expressing that sentiment. Because the signature pages often contained no indication of which cover letter they belonged with, it was not possible to count signatures to double check the accuracy of the signature counts provided in the cover letters. We therefore have had to rely on the petition organizers' signature counts.

Note that some of the "petition cover letters" did not reference petitions, but rather some number of individual comments submitted separately to the Forest Service. For these, we assumed that the individual comments were included either in the set of comments analyzed by the machine-assisted methods described above, or among the postcards evaluated manually.

Results

Individual Comments

There were 81,032 comments reviewed by the machine learning algorithm. Of those results, 1,464 comments, or 1.8% opposed the withdrawal and 79,568 comments, or 98.2% supported the withdrawal. The majority of comments submitted, regardless of for or against the proposal, were form letters.

The results from the machine analysis are divided into four categories:

- A. Opposed to withdrawal and mentions economics.
- B. Opposed to withdrawal for non-economic reasons.
- C. In favor of the withdrawal and mentions economics.
- D. In favor of the withdrawal for non-economic reasons.

Economic concerns mentioned in these the A and C groups include revenue for school funding, average wages, employment opportunities in mining, recreation/tourism, and in other industries, future visitation and visitor spending, job security, in- and out-migration from the Arrowhead region and Minnesota, property value, and future investment opportunities.

Table 2. Sentiment Regarding Mineral Withdrawal and Economic Concern Expressed in Individual Comment Letters (n=81,032)

Category	Percent of Comments
A. Opposed to withdrawal and mentions economics (n=1,219)	1.50%
B. Opposed to withdrawal for non-economic reasons (n= 245)	0.30%
C. In favor of the withdrawal and mentions economics (n=48,965)	60.43%
D. In favor of the withdrawal for non-economic reasons (n= 30,603)	37.77%

Postcards and Petitions

There were 5 petitions, all supporting the withdrawal, and including a total of 39,972 signatures. In addition, we manually reviewed 143 records containing 4,056 comments. Of those, 1,916 comments, or 47.24% opposed the withdrawal and 2,140 (52.76%) supported the withdrawal.

If petition signatures and the postcards and other comments are considered together, 42,112 persons (95.65%) expressed an opinion in favor of the withdrawal, and 1,916 (4.35%) expressed an opinion opposed to the withdrawal.

Table 3. Sentiment Regarding Mineral Withdrawal Expressed in Postcards and Petitions

	Percent of Comments in the same Format	Percent of all Comments
Postcards opposing the withdrawal (n=1,916)	47.24%	4.35% opposed
Postcards supporting the withdrawal (n=2,140)	52.76%	95.65% in favor
Petition signers supporting the withdrawal (n=39,972)	100%	