Modeling of volume change with iron in ferropericlase (Mg,Fe)O

Final Report CMSE - 802
April 21, 2025

By Luisa Chavarria

The repository is stored in GitHub: https://github.com/luisachavarria2/CMSE802_FinalProject.git

1. Introduction

Solid solutions can have different styles of mixing. In the ideal mixing, the volume of the solution changes as the concentration of one of the elements fluctuates. As part of my PhD, I'm studying the solid solution of ferropericlase (Mg,Fe)O. This mineral is the second most abundant mineral in the Earth's lower mantle and has a simple cubic crystallographic structure (Figure 1). My project looks to study the incorporation of sodium inside the structure as a substitution for iron and I've been using three powder samples of (Mg,Fe)O with different concentrations of iron. The samples were provided from a collaborator in Northwestern University. They also supplied an XRD spectra that was used to calculate the atomic volume of the sample.

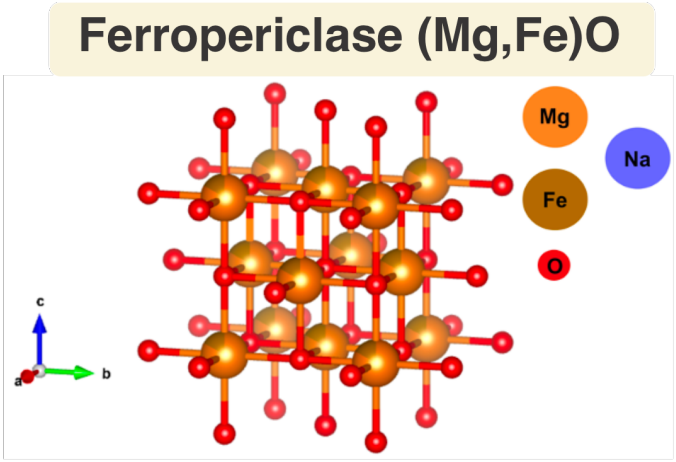

Figure 1. Lattice structure of (Mg,Fe)O with the location of the atoms inside the structure. Sodium (Na) can occupy the positions of Mg or Fe.

Initially, the collaborators provided an estimated composition, however, as part of the characterization, the samples require a verification to double-check that the concentration of iron is correct for the future analysis. Then, this project looks to use techniques of modeling to characterize the concentration of iron in the samples using as starting point the atomic volume calculated from XRD data. To do this, we are going to compare the Vegard's law volume prediction with the linear regression using least squares fit.

The main question of this project is: can a least squares fit model have a better prediction than the Vegard's law? This comparison will allow to determine which approach is better, understand how the unit cell changes as iron is incorporated, and then calculate the concentration of iron of the samples received from Northwestern University based on the XRD volume. Vegard's law is going to be used as a null hypothesis to verify how ideal is the mixing of the samples we received.

Additionally, by studying the substitution of magnesium by iron, it is a first approximation to understand how substitution affects the dimensions of the unit cell of ferropericlase. This approach could be useful in the future for understanding how atoms like Na could be incorporated in the structure. In this case, iron is smaller than sodium, but it can be used as an example of how large atoms change the dimensions of the unit cell. Linear algebra is going to be implemented to calculate an equation that allows the evaluation of the relation between volume and iron concentration.

Objectives:

1. Compare the least squares fit (LSF) vs Vegard's law.
2. Characterize the change in the unit cell as the iron concentration increase
3. Calculate the concentration of the ferropericlase samples from Northwestern University.

2. Methodology

The data that was used for this project was compiled from several references extracted from the database https://rruff.info/index.php and also from relevant papers with XRD data of (Mg,Fe)O at P= 0 GPa and T= 300 K (Table 1). The data was very limited so creating models for predicting the changes of volume according to the iron concentration is relevant to understand this solid solution.

Table 1. Compiled data of (Mg,Fe)O with the iron concentration (%) and the corresponding atomic volume (A^3)

Sample name	Fe (%)	Volume (A^3)	Reference
Fp 25	25	76.42	This Study
Fp 50	50	77.73	This Study
MgO	0	74.99	Hazen (1976)
MgO	0	74.69	Boiocchi et al (2001)
Fp2.6	2.6	74.91	Boiocchi et al (2001)
Fp4.2	4.2	75.03	Boiocchi et al (2001)
Fp17	17	75.85	Matsui et al (2012)

Fp25	25	76.34	Mao et al (2011)
Fp35	35	77.79	Chen et al (2012)
Fp40	40	77.42	Fei et al (1992)
Fp48	48	77.29	Solomatova et al (2016)
Mw60	60	78.43	Richet et al (1988)
Mw80	80	79.26	Richet et al (1988)
Fp84	84	78.32	Wicks (2017)
Fp 94	94	79.23	Wicks (2017)
FeO	100	79.674	Wyckoff and Crittenden (1926)
FeO	100	80.108	Wyckoff (1963)

This project uses linear algebra to model the change of concentration of volume as the iron concentration increases in ferropericlase (Mg,Fe)O. The implementation included Least Square Fit (LSF) and the linear model using Vegard's Law for an overdefined system.

The Vegard's law states that the lattice parameter (or molar volume) of a solid solution varies linearly with the concentration of its constituents.

$$a_{\mathrm{A}_{(1-x)}\mathrm{B}_x} = (1-x)\, a_{\mathrm{A}} + x\, a_{\mathrm{B}}$$

Where a is the lattice parameter, A and B are the constituents of the solid solution, and x is the fraction concentration. Then, applied to the system (Mg,Fe)O, this would be the model that is going to be implemented in this project using the volume (V) instead of the lattice parameter (a):

$$V(x) = (1-x)V_{\mathrm{MgO}} + xV_{\mathrm{FeO}}$$

Where:
V(x) = volume of $(Mg_{1-x}Fe_x)O$
V_{MgO} = volume of end member MgO
V_{FeO} = volume of end member FeO
x = fraction of Fe in the solid solution

Vegard's Law is useful to obtain a rough estimate of the composition of a solution where the experimental data are not available. The expectation is that the solid solution behaves linearly, and the deviations are caused by site ordering, inhomogeneities, and changes in oxidation state.

On the other hand, the Least Square Fit (LSF) has the form y = c1x + c2. Where:

- y is v and it's linear. It represents the volume of ferropericlase
- x is the concentration of iron in ferropericlase

- a, b are the coefficients we are trying to define. Where c1 is the slope of the fitted line and b is the intercept of the line in the y axes

Then, the equation we want to fit is:

$$\mathbf{V} = ax + b$$

Then, the least squared fit and the Vegard's law are expected to be similar for an ideally mixed solution.
The calculation of the iron concentration based on the volume used the equation of volume and solved for x in the LSF section. Then:

$$x = \frac{V-b}{a}$$

To ensure the quality of the fit of the LSF model, we calculated the R^2 and verified that most of the data were inside the 95% band of confidence.

Regarding the code, different methods were used to ensure the optimization and error handling:

- **Optimization techniques implemented**

The method curve_fit from scipy.optimize was used in the module linear_lsf.py to optimize the calculation of the linear model. Additionally, the scripts were divided into modules that allows optimize the running time and an easier implementation and reusability of the codes.

- **Error handling and robustness measures**

Most of the codes include error handling sections were errors are raised if there is an inconsistency or there are exceptions. For example, the data_validator.py file, includes the function check_data_validity. This function evaluates the quality of the data that is uploaded for this project and it raises ValueErrors if x and V are empty, contain NaN values, have different lengths or infinite values.

Validation and Testing

- **Validation methodology**

This is the implementation of the test of the iron calculation in the file fe_calculation.py. Unittest was used for the testing.

- **Test cases and coverage**

Test Case 1. fe_calculation
Test objective: verify that the function estimate_fe estimates the concentration of iron correctly based on the linear equation:

$$x = \frac{V-75.074}{4.816}$$

where volume is the input in A^3 and x is the output (Fe concentration).

Inputs:
There are 8 test cases where the input corresponds to 0%, 50%, and 100% concentration of Fe. Also, there are some cases where there volume is below the minimum volume, there is invalid data like strings, or missing input. Also, values that are very large or small, or are negative.

Expected output:
For example, for the Test case 5 where the input is none, the output is an exception raised where there is a TypeError indicating that the input is incorrect.

Test steps:

1. Prepare the input value
2. Call the function “estimate_fe(volume)
3. Verify the output and check the returned float value
4. Handle the edges and exceptions

Testing implementation:
The script test_fe_calculation.py can be executed in the terminal using the following command:

- python -m unittest test_fe_calculation.py

This is how the test looks like:
The test includes 6 functions:

- test_zero_concentration : zero Fe concentration (MgO)
- test_half_concentration: 50% Fe concentration (Mg,Fe)O
- test_full_concentration: 100% Fe concentration (FeO)
- test_negative_concentration
- test_invalid_input
- test_print_output

- **Quantitative validation results**

As an example, I run the following code:
python -m unittest test_fe_calculation.TestEstimateFe.test_zero_concentration
and obtained the correct answer for the test case where the volume is 75.07 A^3 and then the iron concentration is zero.

- **Analysis of accuracy and reliability**

To analyze the accuracy of the least squared it was compared the LSF vs the Vegard’s Law and it was found that both of them are overlapping. Then, this means that it is reliable to use the LSF model to calculate the concentration of iron and to estimate variations of volume based on iron.

3. Results

The figure 1 shows the initial fit of the end members MgO (Fe= 0) and FeO (Fe= 1). The Vegard's law is a constrained fit where the volume is going to increase as the concentration of iron is higher.

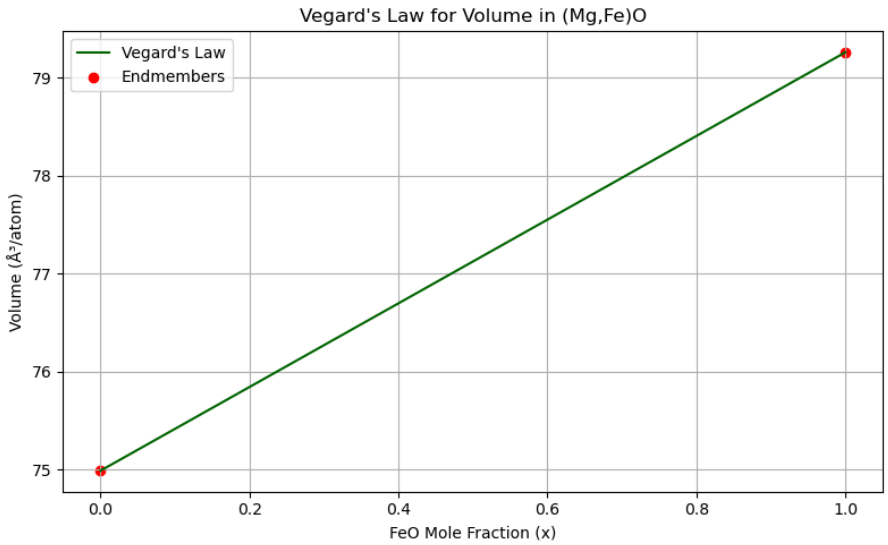

Figure 1. Linear fit of the end members of (Mg,Fe)O based on the Vegard's Law

The second method that was used to understand the change of volume with the iron concentration was the least squares fit where several experimental data from literature was fitted, and the equation of volume was estimated (Figure 2).

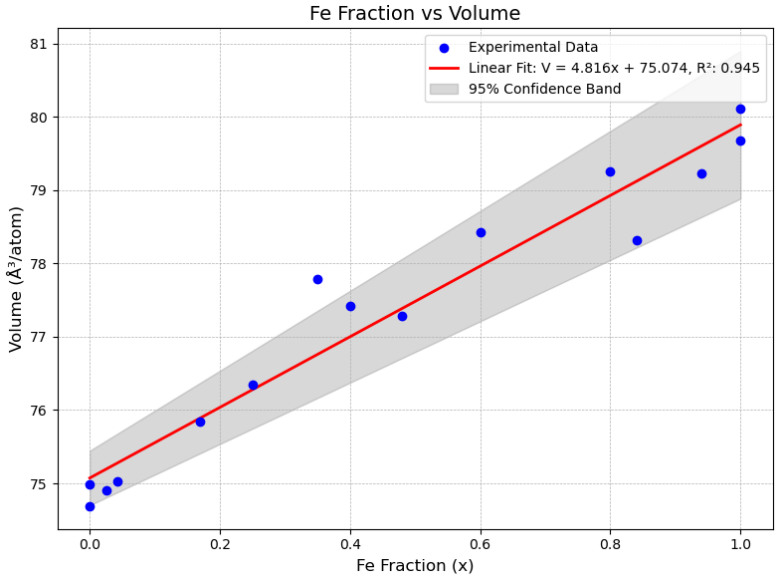

Figure 2. Linear fit of the experimental data using the LSF model

The equation that was obtained was V= 4.816x + 75.074 with a R^2 of 0.945 that shows that the fit is good. In figure 3, it is shown a comparison of the linear fit and the Vegard's law. The deviation from the Vegard's law was estimated to be around 0.286% with a mean absolute deviation of 0.219 A^3/atom. The deviation is mostly positive (Figure 4).

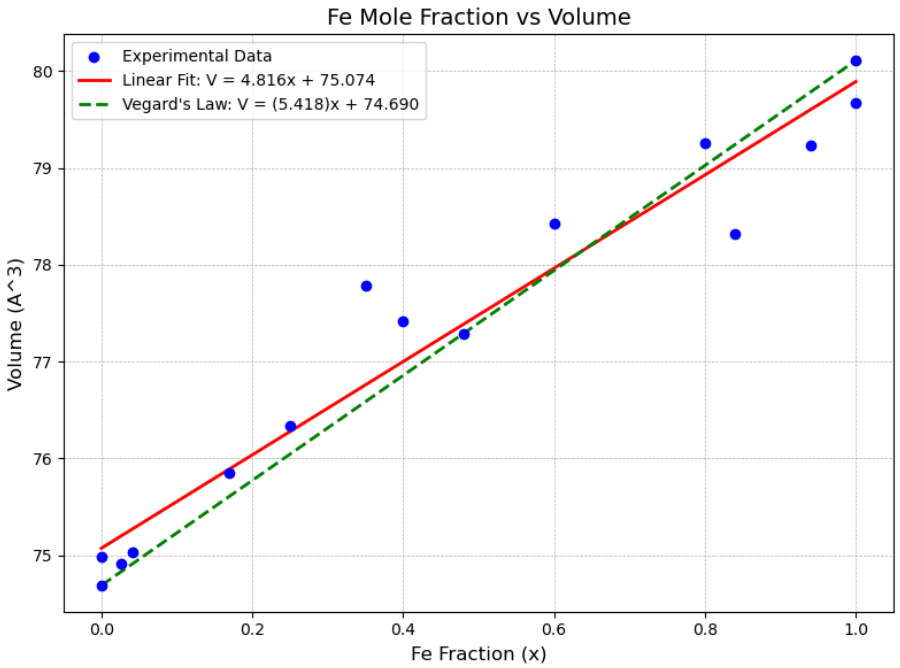

Figure 3. Fe fraction vs volume contrasting the linear fit vs the Vegard's law.

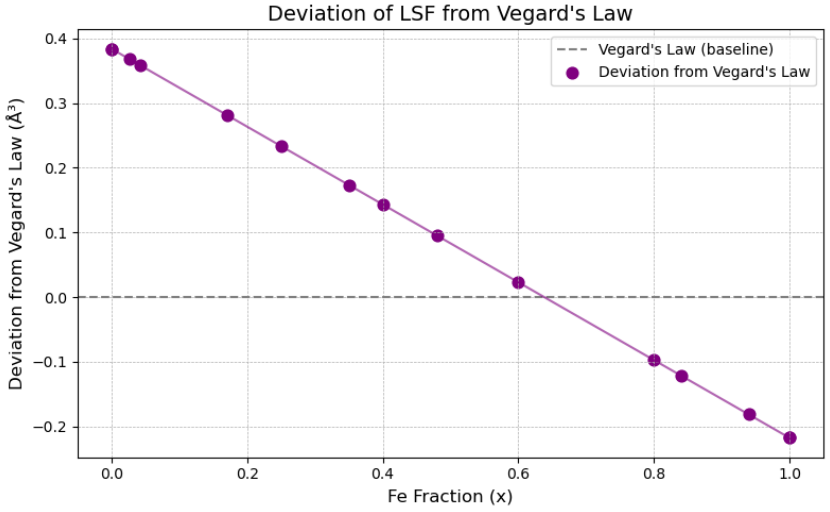

Figure 4. Fe fraction vs deviation from Vegard's law.

Using the equations obtained with the linear fit and the Vegard's law, it was calculated and represented graphically how the unit cell of (Mg,Fe)O changes and expands as the iron concentration increases. The figure 5 shows that the volume for the intermediate compositions (Fe25, Fe50, and Fe75) is higher for the LSF compared to the Vegard's law.

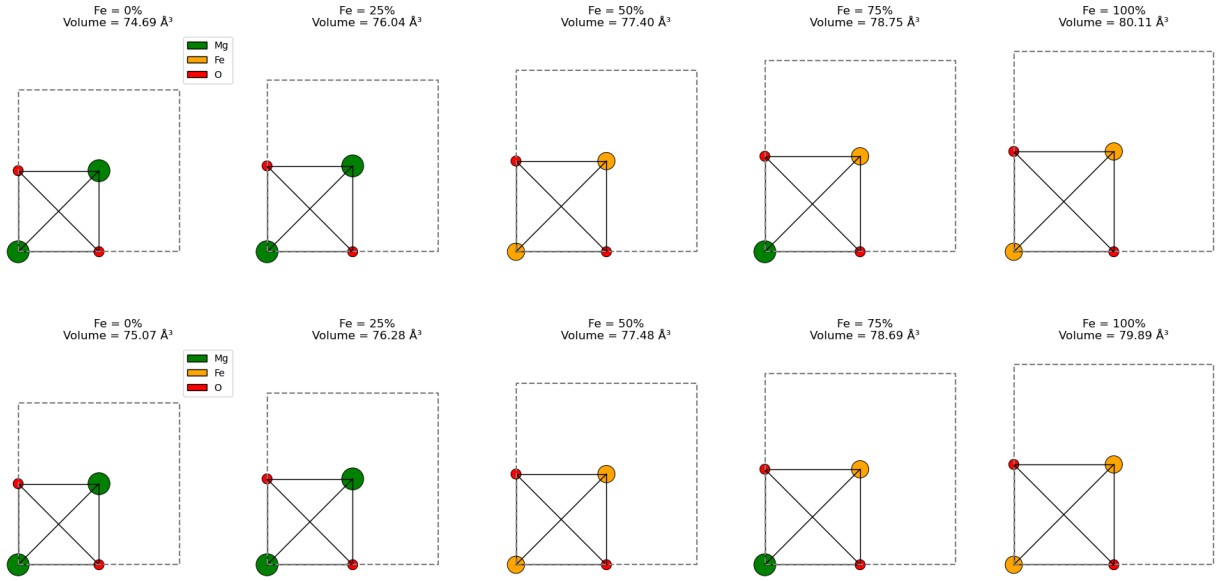

Figure 5. Variation in the unit cell volume as a function of the iron concentration. Comparison between the first row where volume was calculated based on the Vegard's law linear fit and the second row based on the LSF model.

Based on the equations for Vegard's law and the LSF model, the iron concentration was estimated for the powders of Ferropericlase for the samples provided by the collaborators. Table 2 shows a comparison between both models in the estimation of the iron concentration. The results shows that LSF is better for estimating intermediate compositions as the samples 1 and 2 are closer in composition to Fe25 and Fe50 respectively as expected.

Table 2. Comparison of the iron concentration calculated based on the Vegard's law vs the LSF model

Sample	**1 – LSF**	**1-Vegard's law**	**2-LSF**	**2-Vegard's law**
Volume	76.42	76.42	77.73	77.73
Iron concentration	27.95%	31.93%	55.15%	56.11%

4. Synthesis and Discussion

The results showed that the solid solution of ferropericlase (Mg,Fe)O follows an ideal mixing where the volume of the unit cell increases as the concentration of iron is higher. Computational methods showed that the Vegard's law hypothesis was approved however, the LSF method could be more reliable for estimating the volume and iron concentrations of intermediate compositions and not only the end members.

Using both methods, I was able to characterize some samples that belong to my PhD research to evaluate what model would be useful for the characterization of the XRD data that was acquired previously. So, using these models was helpful for the fully characterization of the iron concentration in the samples. Next steps includes to improve the concentration of iron model and also to add some more of my data to evaluate the accuracy of the model

1. References

- Citations for any methods, libraries, or sources used

1. OpenAI. ChatGPT. April 16, 2025. OpenAI. https://chat.openai.com. (Planning and coding of the project)
2. The data from the experimental data was obtained from https://rruff.info/index.php