LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-582974

Analysis of the Formation Enthalpy Dataset

C. Kamath

September 19, 2012

Disclaimer

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Analysis of the Formation Enthalpy Dataset

Chandrika Kamath
kamath2@llnl.gov
Lawrence Livermore National Laboratory
September 14, 2012

Contents

1 Introduction

This report summarizes the work done as part of the *MINDES: Data Mining for Inverse Design* project [3] to mine the datasets generated by the Center for Inverse Design [1], an Energy Frontier Research Center (EFRC) of the Office of Science, US Department of Energy. In the course of the MINDES project, two datasets were analyzed, one on the formation enthalpy of spinels, and the other on the band gap type of the class of ternary compounds generated at NREL; the former is the focus of this report.

The dataset considered in this analysis is from computer simulations of spinel materials. Given a spinel material represented as A2BO4, where A and B are atomic species (O represents oxygen), with the ratio of atoms A:B:O = 16:8:32, simulations are run to obtain the formation enthalpy of a new compound derived from A2BO4 by introducing an impurity in the form of either a vacancy or a substitution in one of the locations of the A, B, or O atoms. The focus of the analysis is to determine, if, given a sample of compounds derived from a spinel material and the corresponding formation enthalpy for these compounds, is it possible to predict the formation enthalpy of a compound derived from the same spinel material using an impurity not considered previously? We expect that the ideas explored in the course of this study will also apply in the analysis of other physical properties of compounds.

This analysis falls in the broad area of design of computer experiments [2], where an ensemble of simulations is used to guide physical experiments and gain insights into the design space which maps the inputs of the simulations to the output(s). As the simulations are often computationally expensive, the ensemble must be carefully designed to gain the greatest insights into the physical phenomenon of interest using as few simulations as possible. A possible solution is to consider an incremental approach where we analyze the input/output data from the simulations that have already been run to identify the next set of simulations such that these new simulations would add the greatest insights, by either refining the original data set in a region of interest, or exploring new regions in the design space.

Therefore, in the context of our problem, we can analyze the simulations that have been run thus far to understand what properties of the elements comprising a compound are relevant to, and predictive of, the formation enthalpy. By combining this information with physics insights, we can suitably create other compounds and expect that their simulation will likely indicate the compounds to have the formation enthalpy predicted by the analysis.

In this report, we start with a description of the data available for analysis in Section 2, and briefly discuss the two types of datasets - the periodic table dataset and the properties dataset - that we can create from the data provided. For the three spinel materials considered in this work - Co2ZnO4, Rh2ZnO4, and Mn2CrO4 - we discuss the analysis results for the periodic table dataset in Section 3 and the properties dataset in Section 4. We conclude with a brief summary and some thoughts for future work.

Disclaimer - This report is written from the point of view of a data miner, not a materials scientist, a physicist, or a chemist. The analysis is purely data driven and is not influenced by any domain-specific biases. Also, any conclusions drawn must be interpreted with care as the analysis reflects the characteristics and quality of the data provided; the availability of additional data may change the results.

name	symbol	group	column	row
mulliken jaffe	single bond radius	pauling	molar volume	atomization
sound velocity	sanderson	atomic weight	critical temperature	triple bond radius
melting point	allen	thermal conductivity	electrical resistivity	vaporization
atomic number	rigidity modulus	covalent radius	fusion	pettifor
bulk modulus	poisson ratio	van der waals radius	electron affinity	boiling point
density	double bond radius	allred rochow	young modulus	thermal expansion
atomic radius	orbital radii s	orbital radii o	first ionization energy	

Table 1: Information available for the different atomic species, including the material properties. The group, row, and column refer to the location of the species in the periodic table. Not all properties are available for all species.

2 Description of the data

There are three types of data available for the analysis described in this report:

- **Formation enthalpy values:** These are available for compounds generated from three spinel materials: Co2ZnO4, Cr2ZnO4, Mn2CrO4. These compounds are created by introducing an impurity, which can be either a vacancy at one of the locations in the spinel structure or a substitution of one atom of an atomic species by another atomic species. These impurities can be at different charge states, with a different formation enthalpy associated with each charge state. For Co2ZnO4, the dataset has 53 compounds, while Cr2ZnO4 has 52 compounds and Mn2CrO4 has 49 compounds.
- **Structure data;** These provide the locations of the different atoms in 3-D space. Some of the structure files are missing as Co2ZnO4 has 32 structure files and Cr2ZnO4 has 47 structure files, though the two materials have 53 and 52 compounds, respectively. There were no files provided for the Mn2CrO4 compounds. The structures files are not used in the present analysis so the missing files are not an issue. However, as observed later in the report, the structure data may be relevant in future work.
- **Material properties:** These are the properties of the different atomic species as listed in Table 1. However, not all properties are available for all atomic species.

2.1 Creating the datasets to describe the compounds

Our goals in this analysis are to identify the atomic properties that are predictive of the formation enthalpy and to determine if we can predict the enthalpy for a new compound. We are also interested in seeing how much of the knowledge gained in the analysis of one spinel material carries over to another.

To accomplish these goals, we need to convert the data available into an appropriate form so that each compound is represented by a series of properties (called "features" in the data mining literature) and its formation enthalpy. Thus, for a spinel material, say Co2ZnO4, each

compound obtained by adding an impurity would represent a row of a table, with the columns representing the properties describing the compound and its formation enthalpy.

Note: In our work, we consider the compounds associated with the three spinel materials as separate datasets. We will refer to the original compound as the "spinel material" or just "material" (e.g., Mn2CrO4) and compounds derived from this as "derived compounds" or just "compounds" (e.g., Co2ZnO4_Mg_on_Co_0 which represents the spinel material Co2ZnO4, with one Co atom replaced by Mg, in charge state 0).

There are two challenges in creating the table for each material. First, how do we define the columns of the table so we can represent both vacancies and substitutions in similar manner, that is, each column in the table represents the same feature for both types of impurities? Second, if we use the material properties (Table 1) to represent each atomic species, how do we suitably modify the properties for a species if one of its atoms is missing or replaced by another? For example, in Co2ZnO4, we have 8 Zn atoms, 16 Co atoms, and 32 O atoms, so a reasonable first attempt is to have three groups of columns, one representing the properties for each atomic species. If a Co atom is replaced by a Ca atom, then we have introduced a fourth species. But, if we replace a Co atom by a Zn atom, we have introduced an impurity in a specific location (a Co atom), without changing the number of atomic species in the compound. Similarly, a vacancy in any one of the three species would reduce the number of atoms for that species, while the number of species remained three. To accommodate all these cases in a consistent manner and to manage the large number of properties representing each atomic species, we created two types of datasets for each spinel material:

- **Periodic table dataset:** In this smaller dataset, we consider just the periodic table information for each species to determine if we can learn any patterns from the data. More details and the results of the analysis are described in Section 3.
- **Properties dataset:** Here, instead of the location of the impurity in the periodic table, we consider the properties associated with the constituent atomic species. The intent is to determine if the values of the properties are reflective of the formation enthalpy. More details and the results of the analysis are described in Section 4.

3 Analysis of the periodic table data

In this dataset, we consider just the periodic table information for each species; this includes five quantities: its symbol, the number of atoms, and the group, row, and column of the periodic table for the species.

To accomodate compounds created by either a vacancy or a substitution in a single table, we consider the table to be formed by a group of columns describing each of the three main species in the spinel (Co, Zn, and O in the case of Co2ZnO4), and a group of columns describing the impurity associated with each species. Since there is only one impurity being considered in our analysis (an atom is either removed or substituted by another), there is some redundancy in this representation. For example, in the Co2ZnO4 data, if a Co atom is replaced by another species, we know that there will not be any impurity associated with the Zn or O species. In such cases, the "impurity" is identified by the sympol "XYZ" and all its periodic table information is set to zero. Further, in our dataset, the only "impurity" associated with the O

Compound ID	natoms Zn	symbol imp_Zn	natoms imp_Zn
PTgroup imp_Zn	PTrow imp_Zn	PTcol imp_Zn	natoms Co
symbol imp_Co	natoms imp_Co	PTgroup imp_Co	PTrow imp_Co
PTcol im_Co	natoms O	charge state	fenthalpy

Table 2: The information in the periodic table (PT) dataset using Co2ZnO4 as the spinel material. The periodic table information includes the group, the row, and the column for the atomic species. "Imp_XX" represents the information for the species which acts as the impurity for XX. If there is no such species, it is represented by the symbol "XYZ" and all the periodic table information is set to zero. The compound ID is a string generated by concatenating the symbols for the spinel material with information on the impurity and the charge state, e.g., Co2ZnO4_Mg_on_Co_0.

atom is a vacancy. This is handled simply by reducing the value of the feature representing the number of O atoms; there are no columns to indicate an impurity associated with the O species.

We also observe that the periodic table information for the three main species (Co, Zn, and O in the case of Co2ZnO4), are the same for all compounds, so we do not need to include the columns for these features in the table. With these simplifications, the periodic table dataset contains the variables listed in Table 2.

Caveat: These observations in the next three sections are based on a relatively small dataset (53 examples for Co2ZnO4, 52 for Rh2ZnO4, and 49 for Mn2CrO4). They are more qualitative in nature and may reflect the set of samples in the data rather than the physics. Therefore, any interpretation and extrapolation must be done with care.

3.1 Analysis of the Co2ZnO4 periodic table dataset

For Co2ZnO4, we first simplified the analysis by assigning a label of "low", "med", or "high" to each compound based on the value of formation enthalpy, which varies from -0.08 to 6.6 for the 53 compounds. The first attempt selected thresholds that split the 53 examples into three equal groups. However, this led to some cases which were close to the threshold, which was then changed so there were no examples near the threshold values. This led to compounds with formation enthaply less than or equal to 1.4 to be labeled "low" (22 examples) and those higher than 3.3 to be labeled as "high" (14 examples) with the remaining 17 examples being labeled "med".

We make the following observations on the data:

1. It appears that the substitutions (for both Co and Zn) are from either column 2 or row 4 of the periodic table.

2. When the compounds are listed in increasing value of f-enthalpy, examples with higher f-enthalpy tend to have negative charge state, while those with lower enthalpy tend to have positive charge states.

3. When the compounds are listed in increasing value of f-enthalpy, examples with lower f-enthalpy (with labels "low" or "med") have either a substitution in Co or a substitution

Compound ID	impurity	PT row	PT col	charge state	f-enthalpy	label
Co2ZnO4_Mg_on_Co_0	Mg	3	2	0	0.700634	low
Co2ZnO4_Mg_on_Co_-1	Mg	3	2	-1	1.005975	low
Co2ZnO4_Be_on_Co_0	Be	2	2	0	1.217644	low
Co2ZnO4_Be_on_Co_-1	Be	2	2	-1	1.321137	low
Co2ZnO4_Ca_on_Co_0	Ca	4	2	0	1.883827	med
Co2ZnO4_Ca_on_Co_-1	Ca	4	2	-1	2.475118	med
Co2ZnO4_Sr_on_Co_0	Sr	5	2	0	3.266471	med
Co2ZnO4_Sr_on_Co_-1	Sr	5	2	-1	3.951836	high
Co2ZnO4_Ba_on_Co_0	Ba	6	2	0	5.334465	high
Co2ZnO4_Ba_on_Co_-1	Ba	6	2	-1	6.102002	high

Table 3: Listed in increasing value of f-enthalpy, the compounds created from Co2ZnO4 by substituting Co with an element from column 2 of the periodic table (PT).

in Zn. The first compound resulting from a vacancy in this ordering of f-enthalpy values occurs at index 27 (out of 53). Note that the converse is not true, that is, a substitution in Co or a substitution in Zn does not result in a lower f-enthalpy.

4. Equivalently, if the impurity is a vacancy in Co, or a vacancy in Zn, or a vacancy in O, then the f-enthalpy is "med" or "high". The converse is not true as "med" or "high" f-enthalpy can result from substitution impurities.

5. There is one compound with negative f-enthalpy, which is a Cr substitution on Co.

6. For a given species which substitutes a Zn atom, the ones with lower charge state have a higher enthalpy (i.e., if we take the compounds where species X replaces a Zn atom and arrange them by decreasing charge state, the order would lead to increasing f-enthalpy values). However, this is not always the case when we consider species which substitute a Co atom.

7. When we consider compounds obtained by substituting Co with an element from column 2 of the periodic table, we observe that the f-enthalpy increases as we go down the column of the periodic table (with the exception of Be, which appears out of order) as shown in Table 3. Note also that, for a species, the lower charge state has higher f-enthalpy.

8. Similarly, when we consider compounds obtained by substituting Zn with an element from column 2 of the periodic table, we observe that the f-enthalpy increases as we go down the column of the periodic table (with the exception of Be, which appears out of order) as shown in Table 4. In this case, the only charge state is zero.

3.2 Analysis of the Rh2ZnO4 periodic table dataset

For Rh2ZnO4, we simplified the analysis by assigning a label of "low", "med", or "high" to each compound based on the value of formation enthalpy, which varies from -2.28 to 12.2 for

Compound ID	impurity	PT row	PT col	charge state	f-enthalpy	label
Co2ZnO4_Mg_on_Zn_0	Mg	3	2	0	0.109360	low
Co2ZnO4_Be_on_Zn_0	Be	2	2	0	0.416459	low
Co2ZnO4_Ca_on_Zn_0	Ca	4	2	0	1.011856	low
Co2ZnO4_Sr_on_Zn_0	Sr	5	2	0	1.989687	med
Co2ZnO4_Ba_on_Zn_0	Ba	6	2	0	3.521112	high

Table 4: Listed in increasing value of f-enthalpy, the compounds created from Co2ZnO4 by substituting Zn with an element from column 2 of the periodic table (PT).

the 52 compounds (there is a large gap between the highest f-enthalpy and the next highest at 6.5). As with Co2ZnO4, the first attempt at selecting thresholds split the 52 examples into three equal groups. However, this led to some cases which were close to the threshold. So, the threshold was changed so there were no examples near the threshold values. This led to compounds with formation enthaply less than or equal to 1.15 to be labeled "low" (19 examples) and those higher than 2.9 to be labeled as "high" (12 examples) with the remaining 21 examples being labeled "med".

We make the following observations on this dataset:

1. It appears that the substitutions (for both Rh and Zn) are from column 2 or row 4 of the periodic table. There are two exceptions - Rh on Zn substitutions at charge state 0 and 1.

2. When the compounds are listed in increasing value of f-enthalpy, examples with higher f-enthalpy tend to have negative charge state, while those with lower enthalpy tend to have positive charge states

3. When the compounds are listed in increasing value of f-enthalpy, examples with lower f-enthalpy (with labels "low" or "med") have either a substitution in Rh or a substitution in Zn. The first compound resulting from a vacancy with this ordering of f-enthalpy values occurs at index 29 (out of 52). Note that the converse is not true, that is, a substitution in Rh or a substitution in Zn does not result in a lower f-enthalpy.

4. Equivalently, if the impurity is a vacancy in Rh, or a vacancy in Zn, or a vacancy in O, then the f-enthalpy is "med" or "high". The converse is not true as "med" or "high" f-enthalpy can result from substitution impurities.

5. There are three compounds with negative f-enthalpy, all resulting from a Cr substitution on Rh or Zn.

6. The following observation made for Co2ZnO4, does not hold good for Rh2ZnO4: "For a given species which substitutes a Zn atom, the ones with lower charge state have a higher enthalpy (i.e., if we take the compounds where species X replaces a Zn atom and arrange them by decreasing charge state, the order would lead to increasing f-enthalpy values). However, this is not always the case when we consider species which substitute a Co atom."

Compound ID	impurity	PT row	PT col	charge state	f-enthalpy	label
Mg_on_Rh_0	Mg	3	2	0	0.1236964375	low
Mg_on_Rh_-1	Mg	3	2	-1	0.4948394375	low
Ca_on_Rh_0	Ca	4	2	0	0.5133934375	low
Ca_on_Rh_-1	Ca	4	2	-1	1.1453564375	low
Sr_on_Rh_0	Sr	5	2	0	1.3901164375	med
Be_on_Rh_0	Be	2	2	0	1.5986374375	med
Be_on_Rh_-1	Be	2	2	-1	1.7308044375	med
Sr_on_Rh_-1	Sr	5	2	-1	2.0913034375	med
Ba_on_Rh_0	Ba	6	2	0	2.8719954375	med
Ba_on_Rh_-1	Ba	6	2	-1	3.5874034375	high

Table 5: Listed in increasing value of f-enthalpy, the compounds created from Rh2ZnO4 by substituting Rh with an element from column 2 of the periodic table (PT).

Compound ID	impurity	PT row	PT col	charge state	f-enthalpy	label
Mg_on_Zn_0	Mg	3	2	0	0.018697	low
Ca_on_Zn_0	Ca	4	2	0	0.306219	low
Sr_on_Zn_0	Sr	5	2	0	0.875269	low
Be_on_Zn_0	Be	2	2	0	1.110115	low
Ba_on_Zn_0	Ba	6	2	0	1.893281	med

Table 6: Listed in increasing value of f-enthalpy, the compounds created from Rh2ZnO4 by substituting Zn with an element from column 2 of the periodic table (PT).

7. When we consider compounds obtained by substituting Rh with an element from column 2 of the periodic table, we observe that the f-enthalpy increases as we go down the column of the periodic table (with the exception of Be which appears out of order) as shown in Table 5. Note that for a species, the lower charge state has higher f-enthalpy.

8. Similarly, when we consider compounds obtained by substituting Zn with an element from column 2 of the periodic table, we observe that the f-enthalpy increases as we go down the column of the periodic table (with the exception of Be which appears out of order) as shown in Table 6. In this case, the only charge state is zero.

3.3 Analysis of the Mn2CrO4 periodic table dataset

For Mn2CrO4, the formation enthalpy values range from -2.08 to 12.36. This dataset was obtained later in the project, and we decided not to split the data into three groups of "low", "medium" and "high" f-enthalpy values as we found that it did not contribute to the analysis. We make the following observations on this dataset:

1. The substitutions for Cr and Mn tend to be from periodic table rows 4 or column 1, though there are several exceptions to this.

2. The dataset has very few examples of positive charge states. Of the 49 samples, 25 are negative, 21 are zero, and only 3 have a positive charge state.

3. The more negative charge states tend to have higher formation enthalpy.

4. When the compounds are listed in increasing order of f-enthalpy, the lower f-enthalpy compounds tend to be a substitution on Mn or Cr. The first compound resulting from a vacancy in this ordering of f-enthalpy values occurs at index 21 (out of 49). Note that the converse is not true, that is, a substitution in Cr or a substitution in Mn does not result in a lower f-enthalpy value.

5. The three lowest f-enthalpy values (two negative, one near zero) are for Mg substitutions on Cr or Mn.

6. Of the six highest f-enthalpy values, five involve substitutions with Na. These are all the compounds in this dataset that were generated using a substitution with Na.

7. It is curious that the highest and lowest f-enthalpy for this spinel material result from substitutions involving elements from row 3 of the periodic table (columns 1 and 2).

8. For a substitution, the lower charge states (negative is lower than zero, which is lower than positive), have higher f-enthalpy, regardless of which species (Cr or Mn) is being substituted.

3.4 Summary of the analysis of periodic table datasets

The analysis of the periodic table datasets for the three spinel materials - Co2ZnO4, Rh2ZnO4, and Mn2CrO4 - indicates that there are some similarities. First, vacancies tend to result in higher formation enthalpy than substitutions. Second, for substitutions, the lower charge states have higher formation enthalpy values.

Our analysis also indicated that Co2ZnO4 and Rh2ZnO4 were more similar to each other than to Mn2CrO4. However, given the small sample sizes of the compounds generated from the three spinel materials, and the resulting biases in sampling the design space, we need to be careful in extrapolating any of the observations made for these three materials to other spinel materials.

4 Analysis of the properties data

In this dataset, instead of the location of the impurity in the periodic table, we consider the properties associated with the atomic species that form the compound derived from the spinel material through either a vacancy or a substitution of one of the species.. The intent is to determine if the values of the properties are reflective of the formation enthalpy.

The first task in the analysis is to derive an appropriate representation for the properties of each compound. However, there are several challenges to this. First, not all properties are available for all atomic species. Second, we need a representation that treats vacancies and

substitutions similarly. This is difficult because in the case of a vacancy, we still have three atomic species representing the compound, while in the case of a substitution, there is a fourth atomic species. We not only need to include the properties of this fourth species, but also indicate which species it is replacing. One approach is to consider a representation where the properties of each of the three atomic species are concatenated into a feature vector to describe the compound. A vacancy or a substitution is then represented by appropriately changing the properties of the species being "vacated" or "substituted".

We explored two ways in which we can incorporate this change, and suitably represent the properties of the three species (denoted by "A", "B", and "C", where "C" represents oxygen):

- **Using differences:** This is similar to the case of ratios, except we take the differences. If there is no vacancy or substitution for one of the three species, A, B, or C, we assign a property its value, if it is available; and N/A, if it is not. If a species has a vacancy, we assign the property a value of 0.0, if it is available; and N/A, if it is not. If a species has a substitution, we take the difference of the values of the property for the species and the impurity, assuming both are available; and N/A, if one or both values are unavailable. Having generated the properties for each compound derived from a spinel material (such as Co2ZnO4), we then remove any properties assigned a value of N/A.
- **Using ratios:** In this case, if there is no vacancy or substitution for one of the three species, A, B, or C, we assign a value of 1.0 to a property, if it is available; and N/A, if it is not. If a species has a vacancy, we assign the property a value of 0.0, if it is available; and N/A, if it is not. If a species has a substitution, we take the ratio of the property of the impurity to the property of the species, assuming both are available; and N/A, if one or both are unavailable. Having generated the properties for each compound derived from a spinel material (such as Co2ZnO4), we then remove any properties assigned a value of N/A.

Note that missing properties are not included in creating the dataset resulting in fewer columns and different properties for each of the three atomic species and their impurity. It may also mean that the properties used to represent different spinel materials might be different.

We also observe that any approach to using the properties of the atomic species, A, B, and C, to represent a compound also implies that we will have cases where two compounds have the same representation for the properties and differ only in the charge state and the f-enthalpy values. Since there appears to be no direct correlation between the charge state and the f-enthalpy, this observation indicates that we might be missing some information that is relevant to predicting the formation enthalpy of a compound.

Our early experiments with the two representations - one based on ratios and the other on differences - using the data for the Co2ZnO4 spinel indicated that there were several issues with the use of differences in the representation of a compound. First, the use of differences could result in negative values for some properties. Second, there can be a large variation in the values of a property. For example, the atomization property for Co has value 426000, which is the value used in the dataset if there is no substitution or vacancy on Co. However, a Co vacancy is represented with a 0.0 value, and a substitution could result in a value as low as -89000. Thus, the use of differences can introduce an artificially large variation in the values of a property. A similar observation can be made for properties such as thermal expansion,

which has a value of 1.3e-05 for Co. Setting the feature value corresponding to this property to 0.0 in case of a Co vacancy artificially introduces a large variation as all other values for this feature are near 1.0e-05. And finally, since the properties of an atomic species differ in value (e.g., the 426000 for atomization versus 1.3e-05 for thermal expansion), we need to normalize each property prior to the analysis. The artificial variation introduced by the use of differences could adversely affect this normalization.

In light of these observations, we consider only the dataset generated using the ratios of properties in our analysis. We refer to these ratios as "features" representing each compound.

Having generated the properties dataset for the three spinel materials, we found there were still challenges to the analysis:

- There are relatively few samples for analysis for each spinel, with around 50 compounds generated for each of Co2ZnO4, Rh2ZnO4, and Mn2Cr04. This makes it difficult to draw strong conclusions as they may reflect the choice of samples rather than any scientific insights.
- Not all properties are available for all atomic species. If a property is unavailable for even one compound in the dataset, it is removed from consideration. This may result in the removal of features that are relevant to the determination of the formation enthalpy.
- As explained earlier, for each spinel material, any impurity that appears in more than one charge state results in compounds that differ in their representation only in the charge state and formation enthalpy; all other features are identical. As there does not appear to be a correlation between the charge state and the formation enthalpy, this indicates that some information necessary for predicting the f-enthalpy is unavailable in our analysis.
- Since there are few sample compounds for each spinel material, it is unclear if the design space has been adequately sampled.
- Several of the features (the ratios of the properties) are correlated. These are removed prior to analysis (see the results for each spinel material for more details).

Given these issues, we were unable to gain any insights into the data using traditional data analysis techniques such as feature selection or regression.

We next wanted to determine if we could learn something from the properties datasets. For example, for each spinel material, if we considered the compounds with the same value of the charge state, would compounds which were near each other in the feature space have similar formation enthalpy values? If this were the case, it would indicate that, for a given charge state, the representation of the compounds is predictive of the formation enthalpy. Note that the converse is not necessarily true as the design space of compounds is not well sampled. For example, two nearby compounds may have quite different values for formation enthalpy if they belong to two groups which are near each other in feature space, but have very different f-enthalpy values. Or, two groups farther away in feature space, could have very similar f-enthalpy values. However, if there is no correlation between the f-enthalpy of a compound and those of (at least) some of its nearest neighbors, then it is unlikely that the representation we have chosen is an appropriate one. We next describe the results of these analysis for the properties data for the three spinel materials.

Feature to keep	Correlated features to remove
ARsingle_bond_radius	ARmolar_volume, ARcovalent_radius, ARdouble_bond_radius, ARatomic_radius, AR_orbital_radius_s, ARorbital_radius_o
ARpauling	ARsanderson, ARallred_rochow, ARpettifor, ARfirst_ionization_energy
ARatomization	ARmelting_point, ARvaporization, ARfusion, ARboiling_point
ARatomic_weight	ARatomic_number
BRsingle_bond_radius	BRmolar_volume, BRcovalent_radius, BRdouble_bond_radius, BRatomic_radius, BR_orbital_radius_s, BRorbital_radius_o
BRpauling	BRsanderson, BRallred_rochow, BRpettifor, BRfirst_ionization_energy
BRatomization	BRmelting_point, BRvaporization, BRfusion, BRboiling_point
BRatomic_weight	BRatomic_number

Table 7: The features to remove because they are correlated to other features in the Co2ZnO4 dataset. The AR and BR prefix indicate the ratios of the property for species A and its impurity and species B and its impurity, respectively.

ARsingle_bond_radius	ARpauling	ARatomization	ARatomic_weight
ARthermal_conductivity	ARrigidity_modulus	ARdensity	ARthermal_expansion
BRsingle_bond_radius	BRpauling	BRatomization	BRatomic_weight
BRthermal_conductivity	BRrigidity_modulus	BRdensity	BRthermal_expansion
CRatomic_weight	cstate		

Table 8: The uncorrelated features used in the analysis of the properties data for Co2ZnO4. The AR and BR prefix indicate the ratios of the property for species A and its impurity and species B and its impurity, respectively.

Caveat: These observations in the next three sections are based on a relatively small dataset (53 examples for Co2ZnO4, 52 for Rh2ZnO4, and 49 for Mn2CrO4). They are more qualitative in nature and may reflect the set of samples in the data rather than the physics. Therefore, any interpretation and extrapolation must be done with care.

4.1 Analysis of the Co2ZnO4 properties dataset

For Co2ZnO4, we first identified the features that were correlated to the others and removed them from consideration. Table 7 lists the features which are correlated; only one of these is considered in the analysis. Table 8 lists the features that remain after the removal of the correlated ones.

The Co2ZnO4 dataset has 21 compounds with charge state 0 and 14 with charge state -1; the remaining charge states have very few compounds and are not considered in the analysis. Once the correlated features have been removed, we consider each of the charge states separately in our analysis to avoid the issue discussed in the previous section.

In Appendix A, for charge states 0 and -1, we present the distance between each compound and its nearest and farthest neighbors, as well as the nearest two compounds, their distances,

and the corresponding formation enthalpy values for these compounds. All distances considered are Euclidean distances.

An analysis of Tables 12 and 14 indicates that some compounds have nearest neighbors that are quite close, while others have their nearest neighbors much farther away. This may indicate possible locations where additional samples might be added.

The analysis of Tables 13 and 15 is somewhat inconclusive as it is unclear if the nearest two neighbors of a compound have formation enthalpies which are close to the formation enthalpy of the compound. However, we found that there is some pattern in the formation enthalpies of compounds with charge state 0 and charge state -1 that were generated using substitutions. Specifically, for some compounds, the difference in enthalpy values between the nearest neighbors with charge state 0 is very similar to the difference in enthalpy values between the nearest neighbors with charge state -1. This may mean that given three of these values, we may be able to predict the fourth. Note that the distances between the nearest neighbors are calculated using only the features derived from the properties of the atomic species; the charge state is not included.

A closer examination of the Co2ZnO4 data indicated that the difference in formation enthalpy between two compounds in charge state 0 is often close to the difference in formation enthalpy between the same two compounds in charge state -1, regardless of whether the two compounds are nearest neighbors or not (that is, the distance between the features is not relevant). A similar statement can also be made about the other charge states, though the number of samples is far fewer to rule out a chance coincidence.

In light of this observation, it might be worth looking into the compounds that do not exhibit this behavior to understand why this is the case.

4.2 Analysis of the Rh2ZnO4 properties dataset

For Rh2ZnO4, we also started by identifying the features that were correlated to the others and removed them from consideration. Table 9 lists the features which are correlated; only one of these is considered in the analysis. This list is slightly different from the correlated variables for Co2ZnO4 as the sanderson variable was unavaiable and the orbital_radius_o was more closely related to the atomic_weight for the species A feature. The list of features that remain after the removal of the correlated ones are identical to those for Co2ZnO4 (see Table 8).

The Rh2ZnO4 dataset has 21 compounds with charge state 0 and 14 with charge state -1; the remaining charge states have very few compounds and are not considered in the analysis. As with Co2ZnO4, we removed the correlated features and analyzed the data for charge states 0 and -1 separately.

In Appendix B, for charge states 0 and -1, we present the distance between each compound and its nearest and farthest neighbors, as well as the nearest two compounds, their distances, and the corresponding formation enthalpy values for these compounds. All distances considered are Euclidean distances. The observations are very similar to those for Co2ZnO4 and are repeated here for completeness.

Tables 17 and 19 indicate that some compounds have nearest neighbors that are quite close, while others have their nearest neighbors much farther away, indicating possible locations where additional samples might be added.

Feature to keep	Correlated features to remove
ARsingle_bond_radius	ARmolar_volume, ARcovalent_radius, ARdouble_bond_radius, ARatomic_radius, AR_orbital_radius_s
ARpauling	ARallred_rochow, ARpettifor, ARfirst_ionization_energy
ARatomization	ARmelting_point, ARvaporization, ARfusion, ARboiling_point
ARatomic_weight	ARatomic_number, AR_orbital_radius_o
BRsingle_bond_radius	BRmolar_volume, BRcovalent_radius, BRdouble_bond_radius, BRatomic_radius, BR_orbital_radius_s, BRorbital_radius_o
BRpauling	BRallred_rochow, BRpettifor, BRfirst_ionization_energy
BRatomization	BRmelting_point, BRvaporization, BRfusion, BRboiling_point
BRatomic_weight	BRatomic_number

Table 9: The features to remove because they are correlated to other features in the Rh2ZnO4 dataset. The AR and BR prefix indicate the ratios of the property for species A and its impurity and species B and its impurity, respectively.

The analysis of Tables 18 and 20 is somewhat inconclusive as it is unclear if the nearest two neighbors of a compound have similar values of formation enthalpies. However, the pattern we had observed in Co2ZnO4 in the formation enthalpies of compounds with charge states 0 and -1 persists for Rh2ZnO4. We found that the difference in formation enthalpy between two compounds in charge state 0 is often close to the difference in formation enthalpy between the same two compounds in charge state -1, regardless of whether the two compounds are nearest neighbors or not (that is, the distance between the features is not relevant). A similar statement can also be made about the other charge states, though the number of samples is far fewer to rule out a chance coincidence. As with Co2ZnO4, it might be worth looking into the compounds that do not exhibit this behavior to understand why this is the case.

4.3 Analysis of the Mn2CrO4 properties dataset

For the Mn2CrO4 dataset, we found that there was a greater correlation between the features correlated to AR(BR)atomization and AR(BR)pauling than observed in the Co2ZnO4 or Rh2ZnO4 data. In Table 10, the features to remove are assigned to the most closely correlated feature to keep. The uncorrelated features used in the analysis are listed in Table 11; note that these are a different set from Co2ZnO4 and Rh2ZnO4 due to the properties available for the atomic species and the features identified as correlated.

The Mn2CrO4 dataset has 21 compounds with charge state 0 and 17 with charge state -1; the remaining charge states have very few compounds and are not considered in the analysis. As before, we remove the correlated features and analyze charge states 0 and -1 separately.

In Appendix C, for charge states 0 and -1, we present the distance between each compound and its nearest and farthest neighbors, as well as the nearest two compounds, their distances, and the corresponding formation enthalpy values for these compounds. All distances considered are Euclidean distances. Our analysis results are very similar to those of the other two compounds and not repeated here.

Feature to keep	Correlated features to remove
ARsingle_bond_radius	ARmolar_volume, ARcovalent_radius, ARdouble_bond_radius, ARatomic_radius, AR_orbital_radius_s
ARpauling	ARallred_rochow, ARpettifor, ARfirst_ionization_energy, ARbulk_modulus
ARatomization	ARsanderson, ARmelting_point, ARvaporization, ARfusion, ARboiling_point
ARatomic_weight	ARatomic_number
BRsingle_bond_radius	BRmolar_volume, BRcovalent_radius, BRdouble_bond_radius, BRatomic_radius, BR_orbital_radius_s, BRorbital_radius_o
BRpauling	BRallred_rochow, BRpettifor, BRfirst_ionization_energy, BRbulk_modulus
BRatomization	BRsanderson, BRmelting_point, BRvaporization, BRfusion, BRboiling_point
BRatomic_weight	BRatomic_number

Table 10: The features to remove because they are correlated to other features in the Mn2CrO4 dataset. The AR and BR prefix indicate the ratios of the property for species A and its impurity and species B and its impurity, respectively.

ARsingle_bond_radius	ARpauling	ARatomization	ARsound_velocity
ARatomic_weight	ARthermal_conductivity	ARdensity	AR_orbital_radius_o
BRsingle_bond_radius	BRpauling	BRatomization	BRsound_velocity
BRatomic_weight	BRthermal_conductivity	BRdensity	BR_orbital_radius_o
CRatomic_weight	cstate		

Table 11: The uncorrelated features used in the analysis of the properties data for Mn2CrO4. The AR and BR prefix indicate the ratios of the property for species A and its impurity and species B and its impurity, respectively.

4.4 Summary of analysis of properties dataset

The analysis of the properties dataset for the three spinel materials is challenging, especially as it is unclear what is the best way to represent a compound in terms of the properties of its constituent atomic species; the sample size is quite small and the design space not adequately sampled; not all properties are available for all atomic species; and we need to represent compounds that differ only in the charge state, but may have very different formation enthalpy values. All this indicates that some information relevant to predicting the formation enthalpy may be missing. The crystal structure of a compound may play a role here, though converting the locations of the different atoms in 3-D space into relevant features is an open question.

We considered two representations of the properties - one based on differences and the other on ratios between a species and the impurity. We analyzed the dataset generated using the ratios as it was better behaved.

Our attempts to find correlations between the features describing a compound and its formation enthalpy were inconclusive. We did find that, if we consider the compounds generated

using a substitution, the difference between the formation enthalpy of two compounds with charge state 0 appears to be close to the difference between the formation enthalpy of the corresponding compounds in charge state -1. This may make it possible to predict one given the other three. The observation extends to other charge states as well, though there are too few examples to draw a conclusion. However, the observation is not universally true and it may be worth investigating the compounds for which this is the case. In addition, the distance to the nearest neighbor for each compound could be used to add additional sample points as appropriate, resulting in a more complete coverage of the design space.

5 Conclusions and future work

In this report, we discussed the analysis of a data set of compounds derived from threee different spinel materials with the intent of determining if it was possible to predict the formation enthalpy of the compounds. There were several challenges to the analysis, including the small number of compounds available for each spinel, the determination of an appropriate representation of a compound in terms of the characteristics of its constituent atomic species, the suitable incorporation of the charge state in the analysis, and the need to treat impurities and vacancies similarly.

Our analysis indicated that traditional data mining techniques, such as feature selection and regression were not much help in the analysis. We obtained some qualitative insights into the compounds and the prediction of the formation enthalpy by considering two datasets - one based on the periodic table and the other based on atomic properties. The analysis of the latter indicated what appears to be a pattern in the differences in formation enthalpy between two compounds with charge state 0 and the same compounds with charge state -1. This may extend to other charge states, though the we did not have enough samples to validate this observation.

Based on the analysis, an obvious next step would be to analyze further the observation on the differences in formation enthalpy between two compounds with different charge states. Increasing the size of the datasets and seeing how many of the qualitative observations remain valid could be another venue of exploration. And finally, there remains the need to identify an appropriate representation for a compound that adequately reflects the formation enthalpy.

6 Acknowledgment

This work was performed as part of the ARRA-funded DOE SciDAC-e project, *MINDES - Data Mining for Inverse Design*, in support of the Center for Inverse Design EFRC. The analysis was done in collaboration with Mayeul d'Avezac from NREL, who provided the data and the domain expertise. The idea of using ratios of properties was motivated by a suggestion by Haowei Peng of NREL who proposed using the differences of properties. The analysis of Co2ZnO4 and Rh2ZnO4 was funded by the SciDAC-e program; the analysis of Mn2CrO4 was unfunded work. I appreciate the support and interest of Mayeul d'Avezac, as well as the EFRC Directors, William Tumas and Alex Zunger.

LLNL-TR-582974: This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

References

[1] Center for Inverse Design web page, 2012. http://www.centerforinversedesign.org/.

[2] Fang, K.-T., Li, R., and Sudjianto, A. *Design and Modeling for Computer Experiments*. Chapman and Hall/CRC Press, Boca Raton, FL, 2005.

[3] MINDES: Data Mining for Inverse Design Project web page, 2012. https://computation.llnl.gov/casc/StarSapphire/MINDES.html.

A Results for the analysis of properties data for Co2ZnO4

Compound ID	nearest neighbor		farthest neighbor	
	compound	distance	compound	distance
Co2Zn_Zn_vacancy_0	Co2ZnO4_O_vacancy_0	3	Co2ZnO4_Ba_on_Co_0	15.4745
Co2ZnO4_Co_vacancy_0	Co2ZnO4_O_vacancy_0	3	Co2ZnO4_Ba_on_Co_0	16.8052
Co2ZnO4_O_vacancy_0	Co2ZnO4_Co_on_Zn_0	1.87673	Co2ZnO4_Ba_on_Co_0	15.2467
Co2ZnO4_Zn_on_Co_0	Co2ZnO4_O_vacancy_0	2.65904	Co2ZnO4_Ba_on_Co_0	14.4937
Co2ZnO4_Co_on_Zn_0	Co2ZnO4_O_vacancy_0	1.87673	Co2ZnO4_Ba_on_Co_0	15.2965
Co2ZnO4_Be_on_Zn_0	Co2ZnO4_Mg_on_Zn_0	5.2123	Co2ZnO4_Ba_on_Co_0	16.8093
Co2ZnO4_Be_on_Co_0	Co2ZnO4_Mg_on_Co_0	5.89104	Co2ZnO4_Ba_on_Co_0	16.9548
Co2ZnO4_Mg_on_Zn_0	Co2ZnO4_Ca_on_Zn_0	4.21729	Co2ZnO4_Ba_on_Co_0	15.9273
Co2ZnO4_Mg_on_Co_0	Co2ZnO4_Zn_on_Co_0	5.10328	Co2ZnO4_Ba_on_Co_0	12.4667
Co2ZnO4_Ca_on_Zn_0	Co2ZnO4_Sr_on_Zn_0	3.72967	Co2ZnO4_Ba_on_Co_0	16.3977
Co2ZnO4_Ca_on_Co_0	Co2ZnO4_Sr_on_Co_0	4.05854	Co2ZnO4_Ba_on_Zn_0	14.301
Co2ZnO4_Sr_on_Zn_0	Co2ZnO4_Ca_on_Zn_0	3.72967	Co2ZnO4_Ba_on_Co_0	16.667
Co2ZnO4_Sr_on_Co_0	Co2ZnO4_Ca_on_Co_0	4.05854	Co2ZnO4_Ba_on_Zn_0	15.3546
Co2ZnO4_Ba_on_Zn_0	Co2ZnO4_Sr_on_Zn_0	3.74677	Co2ZnO4_Ba_on_Co_0	18.0709
Co2ZnO4_Ba_on_Co_0	Co2ZnO4_Sr_on_Co_0	4.14905	Co2ZnO4_Ba_on_Zn_0	18.0709
Co2ZnO4_Ti_on_Zn_0	Co2ZnO4_V_on_Zn_0	1.64389	Co2ZnO4_Ba_on_Co_0	16.082
Co2ZnO4_Ti_on_Co_0	Co2ZnO4_V_on_Co_0	1.428	Co2ZnO4_Ba_on_Co_0	13.8086
Co2ZnO4_V_on_Zn_0	Co2ZnO4_Ti_on_Zn_0	1.64389	Co2ZnO4_Ba_on_Co_0	15.7202
Co2ZnO4_V_on_Co_0	Co2ZnO4_Ti_on_Co_0	1.428	Co2ZnO4_Ba_on_Co_0	14.1144
Co2ZnO4_Cr_on_Zn_0	Co2ZnO4_V_on_Zn_0	3.69606	Co2ZnO4_Ba_on_Co_0	16.0967
Co2ZnO4_Cr_on_Co_0	Co2ZnO4_O_vacancy_0	1.9918	Co2ZnO4_Ba_on_Co_0	15.6155

Table 12: The distances to the nearest and farthest neighbors for the 21 compounds, with charge state 0, obtained from Co2ZnO4.

Compound ID	f-enthalpy	top two nearest neighbors		
		compound	distance	f-enthalpy
Co2ZnO4_Zn_vacancy_0	2.34246	Co2ZnO4_O_vacancy_0	3	3.22481
		Co2ZnO4_Cr_on_Co_0	3.31169	-0.081488
Co2ZnO4_Co_vacancy_0	2.89641	Co2ZnO4_O_vacancy_0	3	3.22481
		Co2ZnO4_Co_on_Zn_0	3.24378	0.273712
Co2ZnO4_O_vacancy_0	3.22481	Co2ZnO4_Co_on_Zn_0	1.87673	0.273712
		Co2ZnO4_Cr_on_Co_0	1.9918	-0.081488
Co2ZnO4_Zn_on_Co_0	1.04821	Co2ZnO4_O_vacancy_0	2.65904	3.22481
		Co2ZnO4_Co_on_Zn_0	2.93132	0.273712
Co2ZnO4_Co_on_Zn_0	0.273712	Co2ZnO4_O_vacancy_0	1.87673	3.22481
		Co2ZnO4_Cr_on_Co_0	2.34295	-0.081488
Co2ZnO4_Be_on_Zn_0	0.416459	Co2ZnO4_Mg_on_Zn_0	5.2123	0.10936
		Co2ZnO4_Co_on_Zn_0	6.90585	0.273712
Co2ZnO4_Be_on_Co_0	1.21764	Co2ZnO4_Mg_on_Co_0	5.89104	0.700634
		Co2ZnO4_Cr_on_Co_0	6.67931	-0.081488
Co2ZnO4_Mg_on_Zn_0	0.10936	Co2ZnO4_Ca_on_Zn_0	4.21729	1.01186
		Co2ZnO4_Co_on_Zn_0	4.45475	0.273712
Co2ZnO4_Mg_on_Co_0	0.700634	Co2ZnO4_Zn_on_Co_0	5.10328	1.04821
		Co2ZnO4_V_on_Co_0	5.8464	1.0342
Co2ZnO4_Ca_on_Zn_0	1.01186	Co2ZnO4_Sr_on_Zn_0	3.72967	1.98969
		Co2ZnO4_Mg_on_Zn_0	4.21729	0.10936
Co2ZnO4_Ca_on_Co_0	1.88383	Co2ZnO4_Sr_on_Co_0	4.05854	3.26647
		Co2ZnO4_Mg_on_Co_0	5.96399	0.700634
Co2ZnO4_Sr_on_Zn_0	1.98969	Co2ZnO4_Ca_on_Zn_0	3.72967	1.01186
		Co2ZnO4_Ba_on_Zn_0	3.74677	3.52111
Co2ZnO4_Sr_on_Co_0	3.26647	Co2ZnO4_Ca_on_Co_0	4.05854	1.88383
		Co2ZnO4_Ba_on_Co_0	4.14905	5.33446
Co2ZnO4_Ba_on_Zn_0	3.52111	Co2ZnO4_Sr_on_Zn_0	3.74677	1.98969
		Co2ZnO4_Ca_on_Zn_0	7.32168	1.01186
Co2ZnO4_Ba_on_Co_0	5.33446	Co2ZnO4_Sr_on_Co_0	4.14905	3.26647
		Co2ZnO4_Ca_on_Co_0	7.98301	1.88383
Co2ZnO4_Ti_on_Zn_0	3.70883	Co2ZnO4_V_on_Zn_0	1.64389	2.91725
		Co2ZnO4_Co_on_Zn_0	4.52005	0.273712
Co2ZnO4_Ti_on_Co_0	1.74786	Co2ZnO4_V_on_Co_0	1.428	1.0342
		Co2ZnO4_Cr_on_Co_0	3.88805	-0.081488
Co2ZnO4_V_on_Zn_0	2.91725	Co2ZnO4_Ti_on_Zn_0	1.64389	3.70883
		Co2ZnO4_Co_on_Zn_0	3.01461	0.273712
Co2ZnO4_V_on_Co_0	1.0342	Co2ZnO4_Ti_on_Co_0	1.428	1.74786
		Co2ZnO4_Cr_on_Co_0	2.62654	-0.081488
Co2ZnO4_Cr_on_Zn_0	2.32195	Co2ZnO4_V_on_Zn_0	3.69606	2.91725
		Co2ZnO4_Co_on_Zn_0	3.85706	0.273712
Co2ZnO4_Cr_on_Co_0	-0.081488	Co2ZnO4_O_vacancy_0	1.9918	3.22481
		Co2ZnO4_Co_on_Zn_0	2.34295	0.273712

Table 13: The distances to the nearest two neighbors, and their f-enthalpy values, for the 21 compounds, with charge state 0, obtained from Co2ZnO4.

Compound ID	nearest neighbor		farthest neighbor	
	compound	distance	compound	distance
Co2ZnO4_Zn_vacancy_-1	Co2ZnO4_Cr_on_Co_-1	3.31169	Co2ZnO4_Ba_on_Co_-1	15.4745
Co2ZnO4_Co_vacancy_-1	Co2ZnO4_Cr_on_Co_-1	3.83354	Co2ZnO4_Ba_on_Co_-1	16.8052
Co2ZnO4_Zn_on_Co_-1	Co2ZnO4_Cr_on_Co_-1	3.31479	Co2ZnO4_Ba_on_Co_-1	14.4937
Co2ZnO4_Be_on_Co_-1	Co2ZnO4_Mg_on_Co_-1	5.89104	Co2ZnO4_Ba_on_Co_-1	16.9548
Co2ZnO4_Mg_on_Co_-1	Co2ZnO4_Zn_on_Co_-1	5.10328	Co2ZnO4_Ba_on_Co_-1	12.4667
Co2ZnO4_Ca_on_Co_-1	Co2ZnO4_Sr_on_Co_-1	4.05854	Co2ZnO4_Co_vacancy_-1	12.1707
Co2ZnO4_Sr_on_Co_-1	Co2ZnO4_Ca_on_Co_-1	4.05854	Co2ZnO4_Co_vacancy_-1	13.5127
Co2ZnO4_Ba_on_Co_-1	Co2ZnO4_Sr_on_Co_-1	4.14905	Co2ZnO4_Be_on_Co_-1	16.9548
Co2ZnO4_Ti_on_Zn_-1	Co2ZnO4_V_on_Zn_-1	1.64389	Co2ZnO4_Ba_on_Co_-1	16.082
Co2ZnO4_Ti_on_Co_-1	Co2ZnO4_V_on_Co_-1	1.428	Co2ZnO4_Ba_on_Co_-1	13.8086
Co2ZnO4_V_on_Zn_-1	Co2ZnO4_Ti_on_Zn_-1	1.64389	Co2ZnO4_Ba_on_Co_-1	15.7202
Co2ZnO4_V_on_Co_-1	Co2ZnO4_Ti_on_Co_-1	1.428	Co2ZnO4_Ba_on_Co_-1	14.1144
Co2ZnO4_Cr_on_Zn_-1	Co2ZnO4_V_on_Zn_-1	3.69606	Co2ZnO4_Ba_on_Co_-1	16.0967
Co2ZnO4_Cr_on_Co_-1	Co2ZnO4_V_on_Co_-1	2.62654	Co2ZnO4_Ba_on_Co_-1	15.6155

Table 14: The distances to the nearest and farthest neighbors for the 14 compounds, with charge state -1, obtained from Co2ZnO4.

Compound ID	f-enthalpy	top two nearest neighbors		
		compound	distance	f-enthalpy
Co2ZnO4_Zn_vacancy_-1	2.73757	Co2ZnO4_Cr_on_Co_-1	3.31169	2.0137
		Co2ZnO4_V_on_Co_-1	3.73162	3.10042
Co2ZnO4_Co_vacancy_-1	3.58	Co2ZnO4_Cr_on_Co_-1	3.83354	2.0137
		Co2ZnO4_Zn_vacancy_-1	4	2.73757
Co2ZnO4_Zn_on_Co_-1	1.38399	Co2ZnO4_Cr_on_Co_-1	3.31479	2.0137
		Co2ZnO4_V_on_Co_-1	3.60493	3.10042
Co2ZnO4_Be_on_Co_-1	1.32114	Co2ZnO4_Mg_on_Co_-1	5.89104	1.00598
		Co2ZnO4_Cr_on_Co_-1	6.67931	2.0137
Co2ZnO4_Mg_on_Co_-1	1.00598	Co2ZnO4_Zn_on_Co_-1	5.10328	1.38399
		Co2ZnO4_V_on_Co_-1	5.8464	3.10042
Co2ZnO4_Ca_on_Co_-1	2.47512	Co2ZnO4_Sr_on_Co_-1	4.05854	3.95184
		Co2ZnO4_Mg_on_Co_-1	5.96399	1.00598
Co2ZnO4_Sr_on_Co_-1	3.95184	Co2ZnO4_Ca_on_Co_-1	4.05854	2.47512
		Co2ZnO4_Ba_on_Co_-1	4.14905	6.102
Co2ZnO4_Ba_on_Co_-1	6.102	Co2ZnO4_Sr_on_Co_-1	4.14905	3.95184
		Co2ZnO4_Ca_on_Co_-1	7.98301	2.47512
Co2ZnO4_Ti_on_Zn_-1	5.46123	Co2ZnO4_V_on_Zn_-1	1.64389	4.83029
		Co2ZnO4_Cr_on_Zn_-1	5.02217	4.07391
Co2ZnO4_Ti_on_Co_-1	3.98017	Co2ZnO4_V_on_Co_-1	1.428	3.10042
		Co2ZnO4_Cr_on_Co_-1	3.88805	2.0137
Co2ZnO4_V_on_Zn_-1	4.83029	Co2ZnO4_Ti_on_Zn_-1	1.64389	5.46123
		Co2ZnO4_Cr_on_Zn_-1	3.69606	4.07391
Co2ZnO4_V_on_Co_-1	3.10042	Co2ZnO4_Ti_on_Co_-1	1.428	3.98017
		Co2ZnO4_Cr_on_Co_-1	2.62654	2.0137
Co2ZnO4_Cr_on_Zn_-1	4.07391	Co2ZnO4_V_on_Zn_-1	3.69606	4.83029
		Co2ZnO4_Ti_on_Zn_-1	5.02217	5.46123
Co2ZnO4_Cr_on_Co_-1	2.0137	Co2ZnO4_V_on_Co_-1	2.62654	3.10042
		Co2ZnO4_Zn_vacancy_-1	3.31169	2.73757

Table 15: The distances to the nearest two neighbors, and their f-enthalpy values, for the 14 compounds, with charge state -1, obtained from Co2ZnO4.

Compound ID	f-enthalpy	nearest neighbor			
		compound	distance	f-enthalpy	difference
Co2ZnO4_Be_on_Co_0	1.21764	Co2ZnO4_Mg_on_Co_0	5.89104	0.700634	0.517006
Co2ZnO4_Be_on_Co_-1	1.32114	Co2ZnO4_Mg_on_Co_-1	5.89104	1.00598	0.31516
Co2ZnO4_Mg_on_Co_0	0.700634	Co2ZnO4_Zn_on_Co_0	5.10328	1.04821	-0.347576
Co2ZnO4_Mg_on_Co_-1	1.00598	Co2ZnO4_Zn_on_Co_-1	5.10328	1.38399	-0.37801
Co2ZnO4_Sr_on_Co_0	3.26647	Co2ZnO4_Ca_on_Co_0	4.05854	1.88383	1.38264
Co2ZnO4_Sr_on_Co_-1	3.95184	Co2ZnO4_Ca_on_Co_-1	4.05854	2.47512	1.47672
Co2ZnO4_Ba_on_Co_0	5.33446	Co2ZnO4_Sr_on_Co_0	4.14905	3.26647	2.06799
Co2ZnO4_Ba_on_Co_-1	6.102	Co2ZnO4_Sr_on_Co_-1	4.14905	3.95184	2.15016
Co2ZnO4_Ti_on_Zn_0	3.70883	Co2ZnO4_V_on_Zn_0	1.64389	2.91725	0.79158
Co2ZnO4_Ti_on_Zn_-1	5.46123	Co2ZnO4_V_on_Zn_-1	1.64389	4.83029	0.63094
Co2ZnO4_Ti_on_Co_0	1.74786	Co2ZnO4_V_on_Co_0	1.428	1.0342	0.71366
Co2ZnO4_Ti_on_Co_-1	3.98017	Co2ZnO4_V_on_Co_-1	1.428	3.10042	0.87975
Co2ZnO4_Cr_on_Zn_0	2.32195	Co2ZnO4_V_on_Zn_0	3.69606	2.91725	-0.5953
Co2ZnO4_Cr_on_Zn_-1	4.07391	Co2ZnO4_V_on_Zn_-1	3.69606	4.83029	-0.75638

Table 16: The difference between f-enthalpy values for a compound and its nearest neighbor, both with charge states 0, and the corresponding compounds with charge state -1 (which are also nearest neighbors). Note the similarity in the differences, indicating that if three of the values are known, the fourth could be predicted. Compounds obtained from Co2ZnO4.

B Results for the analysis of properties data for Rh2ZnO4

Compound ID	nearest neighbor		farthest neighbor	
	compound	distance	compound	distance
Rh2ZnO4_Zn_vacancy_0	Rh2ZnO4_O_vacancy_0	3	Rh2ZnO4_Ba_on_Rh_0	30.8605
Rh2ZnO4_Rh_vacancy_0	Rh2ZnO4_O_vacancy_0	3	Rh2ZnO4_Ba_on_Rh_0	32.1894
Rh2ZnO4_Zn_on_Rh_0	Rh2ZnO4_Cr_on_Rh_0	3.88682	Rh2ZnO4_Ba_on_Rh_0	28.1569
Rh2ZnO4_Rh_on_Zn_0	Rh2ZnO4_Cr_on_Zn_0	2.65296	Rh2ZnO4_Ba_on_Rh_0	30.8722
Rh2ZnO4_O_vacancy_0	Rh2ZnO4_Cr_on_Rh_0	1.927	Rh2ZnO4_Ba_on_Rh_0	30.7468
Rh2ZnO4_Be_on_Zn_0	Rh2ZnO4_Mg_on_Zn_0	5.2123	Rh2ZnO4_Ba_on_Rh_0	31.5509
Rh2ZnO4_Be_on_Rh_0	Rh2ZnO4_Mg_on_Rh_0	10.7497	Rh2ZnO4_Ba_on_Rh_0	32.4534
Rh2ZnO4_Mg_on_Zn_0	Rh2ZnO4_Ca_on_Zn_0	4.21729	Rh2ZnO4_Ba_on_Rh_0	31.09
Rh2ZnO4_Mg_on_Rh_0	Rh2ZnO4_Zn_on_Rh_0	8.11697	Rh2ZnO4_Ba_on_Rh_0	23.5869
Rh2ZnO4_Ca_on_Zn_0	Rh2ZnO4_Sr_on_Zn_0	3.72967	Rh2ZnO4_Ba_on_Rh_0	31.3335
Rh2ZnO4_Ca_on_Rh_0	Rh2ZnO4_Sr_on_Rh_0	6.63866	Rh2ZnO4_Ba_on_Zn_0	22.9136
Rh2ZnO4_Sr_on_Zn_0	Rh2ZnO4_Ca_on_Zn_0	3.72967	Rh2ZnO4_Ba_on_Rh_0	31.4753
Rh2ZnO4_Sr_on_Rh_0	Rh2ZnO4_Ca_on_Rh_0	6.63866	Rh2ZnO4_Ba_on_Zn_0	26.1638
Rh2ZnO4_Ba_on_Zn_0	Rh2ZnO4_Sr_on_Zn_0	3.74677	Rh2ZnO4_Ba_on_Rh_0	32.2407
Rh2ZnO4_Ba_on_Rh_0	Rh2ZnO4_Sr_on_Rh_0	7.37422	Rh2ZnO4_Be_on_Rh_0	32.4534
Rh2ZnO4_Ti_on_Zn_0	Rh2ZnO4_V_on_Zn_0	1.64389	Rh2ZnO4_Ba_on_Rh_0	31.1695
Rh2ZnO4_Ti_on_Rh_0	Rh2ZnO4_V_on_Rh_0	2.12727	Rh2ZnO4_Ba_on_Rh_0	27.3853
Rh2ZnO4_V_on_Zn_0	Rh2ZnO4_Ti_on_Zn_0	1.64389	Rh2ZnO4_Ba_on_Rh_0	30.9844
Rh2ZnO4_V_on_Rh_0	Rh2ZnO4_Ti_on_Rh_0	2.12727	Rh2ZnO4_Ba_on_Rh_0	27.8151
Rh2ZnO4_Cr_on_Zn_0	Rh2ZnO4_Rh_on_Zn_0	2.65296	Rh2ZnO4_Ba_on_Rh_0	31.1771
Rh2ZnO4_Cr_on_Rh_0	Rh2ZnO4_O_vacancy_0	1.927	Rh2ZnO4_Ba_on_Rh_0	30.2496

Table 17: The distances to the nearest and farthest neighbors for the 21 compounds, with charge state 0, obtained from Rh2ZnO4.

Compound ID	f-enthalpy	top two nearest neighbors		
		compound	distance	f-enthalpy
Rh2ZnO4_Zn_vacancy_0	1.91477	Rh2ZnO4_O_vacancy_0	3	2.83997
		Rh2ZnO4_Cr_on_Rh_0	3.27312	-2.28575
Rh2ZnO4_Rh_vacancy_0	2.08431	Rh2ZnO4_O_vacancy_0	3	2.83997
		Rh2ZnO4_Zn_vacancy_0	4	1.91477
Rh2ZnO4_Zn_on_Rh_0	0.531288	Rh2ZnO4_Cr_on_Rh_0	3.88682	-2.28575
		Rh2ZnO4_O_vacancy_0	4.39799	2.83997
Rh2ZnO4_Rh_on_Zn_0	1.74137	Rh2ZnO4_Cr_on_Zn_0	2.65296	0.231912
		Rh2ZnO4_O_vacancy_0	3.11859	2.83997
Rh2ZnO4_O_vacancy_0	2.83997	Rh2ZnO4_Cr_on_Rh_0	1.927	-2.28575
		Rh2ZnO4_Zn_vacancy_0	3	1.91477
Rh2ZnO4_Be_on_Zn_0	1.11011	Rh2ZnO4_Mg_on_Zn_0	5.2123	0.018697
		Rh2ZnO4_O_vacancy_0	7.21738	2.83997
Rh2ZnO4_Be_on_Rh_0	1.59864	Rh2ZnO4_Mg_on_Rh_0	10.7497	0.123696
		Rh2ZnO4_Cr_on_Rh_0	10.7591	-2.28575
Rh2ZnO4_Mg_on_Zn_0	0.018697	Rh2ZnO4_Ca_on_Zn_0	4.21729	0.306219
		Rh2ZnO4_Rh_on_Zn_0	4.7415	1.74137
Rh2ZnO4_Mg_on_Rh_0	0.123696	Rh2ZnO4_Zn_on_Rh_0	8.11697	0.531288
		Rh2ZnO4_V_on_Rh_0	9.25622	0.778359
Rh2ZnO4_Ca_on_Zn_0	0.306219	Rh2ZnO4_Sr_on_Zn_0	3.72967	0.875269
		Rh2ZnO4_Mg_on_Zn_0	4.21729	0.018697
Rh2ZnO4_Ca_on_Rh_0	0.513393	Rh2ZnO4_Sr_on_Rh_0	6.63866	1.39012
		Rh2ZnO4_Mg_on_Rh_0	11.6526	0.123696
Rh2ZnO4_Sr_on_Zn_0	0.875269	Rh2ZnO4_Ca_on_Zn_0	3.72967	0.306219
		Rh2ZnO4_Ba_on_Zn_0	3.74677	1.89328
Rh2ZnO4_Sr_on_Rh_0	1.39012	Rh2ZnO4_Ca_on_Rh_0	6.63866	0.513393
		Rh2ZnO4_Ba_on_Rh_0	7.37422	2.872
Rh2ZnO4_Ba_on_Zn_0	1.89328	Rh2ZnO4_Sr_on_Zn_0	3.74677	0.875269
		Rh2ZnO4_Ca_on_Zn_0	7.32168	0.306219
Rh2ZnO4_Ba_on_Rh_0	2.872	Rh2ZnO4_Sr_on_Rh_0	7.37422	1.39012
		Rh2ZnO4_Ca_on_Rh_0	13.7106	0.513393
Rh2ZnO4_Ti_on_Zn_0	4.15957	Rh2ZnO4_V_on_Zn_0	1.64389	3.51051
		Rh2ZnO4_Rh_on_Zn_0	4.88414	1.74137
Rh2ZnO4_Ti_on_Rh_0	1.59054	Rh2ZnO4_V_on_Rh_0	2.12727	0.778359
		Rh2ZnO4_Cr_on_Rh_0	5.77581	-2.28575
Rh2ZnO4_V_on_Zn_0	3.51051	Rh2ZnO4_Ti_on_Zn_0	1.64389	4.15957
		Rh2ZnO4_Rh_on_Zn_0	3.27032	1.74137
Rh2ZnO4_V_on_Rh_0	0.778359	Rh2ZnO4_Ti_on_Rh_0	2.12727	1.59054
		Rh2ZnO4_Cr_on_Rh_0	3.84245	-2.28575
Rh2ZnO4_Cr_on_Zn_0	0.231912	Rh2ZnO4_Rh_on_Zn_0	2.65296	1.74137
		Rh2ZnO4_V_on_Zn_0	3.69606	3.51051
Rh2ZnO4_Cr_on_Rh_0	-2.28575	Rh2ZnO4_O_vacancy_0	1.927	2.83997
		Rh2ZnO4_Zn_vacancy_0	3.27312	1.91477

Table 18: The distances to the nearest two neighbors, and their f-enthalpy values, for the 21 compounds, with charge state 0, obtained from Rh2ZnO4.

Compound ID	nearest neighbor		farthest neighbor	
	compound	distance	compound	distance
Rh2ZnO4_Zn_vacancy_-1	Rh2ZnO4_Cr_on_Rh_-1	3.27312	Rh2ZnO4_Ba_on_Rh_-1	30.8605
Rh2ZnO4_Rh_vacancy_-1	Rh2ZnO4_Zn_vacancy_-1	4	Rh2ZnO4_Ba_on_Rh_-1	32.1894
Rh2ZnO4_Zn_on_Rh_-1	Rh2ZnO4_Cr_on_Rh_-1	3.88682	Rh2ZnO4_Ba_on_Rh_-1	28.1569
Rh2ZnO4_Be_on_Rh_-1	Rh2ZnO4_Mg_on_Rh_-1	10.7497	Rh2ZnO4_Ba_on_Rh_-1	32.4534
Rh2ZnO4_Mg_on_Rh_-1	Rh2ZnO4_Zn_on_Rh_-1	8.11697	Rh2ZnO4_Ba_on_Rh_-1	23.5869
Rh2ZnO4_Ca_on_Rh_-1	Rh2ZnO4_Sr_on_Rh_-1	6.63866	Rh2ZnO4_Rh_vacancy_-1	22.32
Rh2ZnO4_Sr_on_Rh_-1	Rh2ZnO4_Ca_on_Rh_-1	6.63866	Rh2ZnO4_Be_on_Rh_-1	25.9883
Rh2ZnO4_Ba_on_Rh_-1	Rh2ZnO4_Sr_on_Rh_-1	7.37422	Rh2ZnO4_Be_on_Rh_-1	32.4534
Rh2ZnO4_Ti_on_Zn_-1	Rh2ZnO4_V_on_Zn_-1	1.64389	Rh2ZnO4_Ba_on_Rh_-1	31.1695
Rh2ZnO4_Ti_on_Rh_-1	Rh2ZnO4_V_on_Rh_-1	2.12727	Rh2ZnO4_Ba_on_Rh_-1	27.3853
Rh2ZnO4_V_on_Zn_-1	Rh2ZnO4_Ti_on_Zn_-1	1.64389	Rh2ZnO4_Ba_on_Rh_-1	30.9844
Rh2ZnO4_V_on_Rh_-1	Rh2ZnO4_Ti_on_Rh_-1	2.12727	Rh2ZnO4_Ba_on_Rh_-1	27.8151
Rh2ZnO4_Cr_on_Zn_-1	Rh2ZnO4_V_on_Zn_-1	3.69606	Rh2ZnO4_Ba_on_Rh_-1	31.1771
Rh2ZnO4_Cr_on_Rh_-1	Rh2ZnO4_Zn_vacancy_-1	3.27312	Rh2ZnO4_Ba_on_Rh_-1	30.2496

Table 19: The distances to the nearest and farthest neighbors for the 14 compounds, with charge state -1, obtained from Rh2ZnO4.

Compound ID	f-enthalpy	top two nearest neighbors		
		compound	distance	f-enthalpy
Rh2ZnO4_Zn_vacancy_-1	12.2039	Rh2ZnO4_Cr_on_Rh_-1	3.27312	0.209424
		Rh2ZnO4_Rh_vacancy_-1	4	2.38295
Rh2ZnO4_Rh_vacancy_-1	2.38295	Rh2ZnO4_Zn_vacancy_-1	4	12.2039
		Rh2ZnO4_Cr_on_Rh_-1	4.34778	0.209424
Rh2ZnO4_Zn_on_Rh_-1	0.694745	Rh2ZnO4_Cr_on_Rh_-1	3.88682	0.209424
		Rh2ZnO4_V_on_Rh_-1	4.87651	3.13748
Rh2ZnO4_Be_on_Rh_-1	1.7308	Rh2ZnO4_Mg_on_Rh_-1	10.7497	0.494839
		Rh2ZnO4_Cr_on_Rh_-1	10.7591	0.209424
Rh2ZnO4_Mg_on_Rh_-1	0.494839	Rh2ZnO4_Zn_on_Rh_-1	8.11697	0.694745
		Rh2ZnO4_V_on_Rh_-1	9.25622	3.13748
Rh2ZnO4_Ca_on_Rh_-1	1.14536	Rh2ZnO4_Sr_on_Rh_-1	6.63866	2.0913
		Rh2ZnO4_Mg_on_Rh_-1	11.6526	0.494839
Rh2ZnO4_Sr_on_Rh_-1	2.0913	Rh2ZnO4_Ca_on_Rh_-1	6.63866	1.14536
		Rh2ZnO4_Ba_on_Rh_-1	7.37422	3.5874
Rh2ZnO4_Ba_on_Rh_-1	3.5874	Rh2ZnO4_Sr_on_Rh_-1	7.37422	2.0913
		Rh2ZnO4_Ca_on_Rh_-1	13.7106	1.14536
Rh2ZnO4_Ti_on_Zn_-1	6.5012	Rh2ZnO4_V_on_Zn_-1	1.64389	5.61178
		Rh2ZnO4_Cr_on_Zn_-1	5.02217	2.58894
Rh2ZnO4_Ti_on_Rh_-1	4.40298	Rh2ZnO4_V_on_Rh_-1	2.12727	3.13748
		Rh2ZnO4_Cr_on_Rh_-1	5.77581	0.209424
Rh2ZnO4_V_on_Zn_-1	5.61178	Rh2ZnO4_Ti_on_Zn_-1	1.64389	6.5012
		Rh2ZnO4_Cr_on_Zn_-1	3.69606	2.58894
Rh2ZnO4_V_on_Rh_-1	3.13748	Rh2ZnO4_Ti_on_Rh_-1	2.12727	4.40298
		Rh2ZnO4_Cr_on_Rh_-1	3.84245	0.209424
Rh2ZnO4_Cr_on_Zn_-1	2.58894	Rh2ZnO4_V_on_Zn_-1	3.69606	5.61178
		Rh2ZnO4_Ti_on_Zn_-1	5.02217	6.5012
Rh2ZnO4_Cr_on_Rh_-1	0.209424	Rh2ZnO4_Zn_vacancy_-1	3.27312	12.2039
		Rh2ZnO4_V_on_Rh_-1	3.84245	3.13748

Table 20: The distances to the nearest two neighbors, and their f-enthalpy values, for the 14 compounds, with charge state -1, obtained from Rh2ZnO4.

Compound ID	f-enthalpy	nearest neighbor			
		compound	distance	f-enthalpy	difference
Rh2ZnO4_Zn_on_Rh_0	0.531288	Rh2ZnO4_Cr_on_Rh_0	3.88682	-2.28575	2.817038
Rh2ZnO4_Zn_on_Rh_-1	0.694745	Rh2ZnO4_Cr_on_Rh_-1	3.88682	0.209424	0.485321
Rh2ZnO4_Be_on_Rh_0	1.59864	Rh2ZnO4_Mg_on_Rh_0	10.7497	0.123696	1.474944
Rh2ZnO4_Be_on_Rh_-1	1.7308	Rh2ZnO4_Mg_on_Rh_-1	10.7497	0.494839	1.235961
Rh2ZnO4_Mg_on_Rh_0	0.123696	Rh2ZnO4_Zn_on_Rh_0	8.11697	0.531288	-0.407592
Rh2ZnO4_Mg_on_Rh_-1	0.494839	Rh2ZnO4_Zn_on_Rh_-1	8.11697	0.694745	-0.199906
Rh2ZnO4_Ca_on_Rh_0	0.513393	Rh2ZnO4_Sr_on_Rh_0	6.63866	1.39012	-0.876727
Rh2ZnO4_Ca_on_Rh_-1	1.14536	Rh2ZnO4_Sr_on_Rh_-1	6.63866	2.0913	-0.94594
Rh2ZnO4_Ti_on_Zn_0	4.15957	Rh2ZnO4_V_on_Zn_0	1.64389	3.51051	0.64906
Rh2ZnO4_Ti_on_Zn_-1	6.5012	Rh2ZnO4_V_on_Zn_-1	1.64389	5.61178	0.88942
Rh2ZnO4_Ti_on_Rh_0	1.59054	Rh2ZnO4_V_on_Rh_0	2.12727	0.778359	0.812181
Rh2ZnO4_Ti_on_Rh_-1	4.40298	Rh2ZnO4_V_on_Rh_-1	2.12727	3.13748	1.2655

Table 21: The difference between f-enthalpy values for a compound and its nearest neighbor, both with charge states 0, and the corresponding compounds with charge state -1 (which are also nearest neighbors). Note the similarity in the differences, indicating that if three of the values are known, the fourth could be predicted. Compounds obtained from Rh2ZnO4.

C Results for the analysis of properties data for Mn2CrO4

Compound ID	nearest neighbor		farthest neighbor	
	compound	distance	compound	distance
Mn2CrO4_Mn_on_Cr_0	Mn2CrO4_Cr_on_Mn_0	11.1046	Mn2CrO4_Li_on_Mn_0	18.4639
Mn2CrO4_Ag_on_Mn_0	Mn2CrO4_Cu_on_Mn_0	0.701159	Mn2CrO4_Li_on_Mn_0	15.2968
Mn2CrO4_Cu_on_Mn_0	Mn2CrO4_Cr_on_Mn_0	0.692069	Mn2CrO4_Li_on_Mn_0	14.988
Mn2CrO4_Na_on_Cr_0	Mn2CrO4_K_on_Cr_0	1.95465	Mn2CrO4_Li_on_Mn_0	16.4218
Mn2CrO4_Li_on_Cr_0	Mn2CrO4_Na_on_Cr_0	8.19348	Mn2CrO4_Li_on_Mn_0	20.4275
Mn2CrO4_Mn_vacancy_0	Mn2CrO4_Cr_on_Mn_0	2.46979	Mn2CrO4_Li_on_Mn_0	16.3904
Mn2CrO4_Zn_on_Cr_0	Mn2CrO4_Ag_on_Cr_0	1.93871	Mn2CrO4_Li_on_Mn_0	14.9409
Mn2CrO4_Ag_on_Cr_0	Mn2CrO4_Cu_on_Cr_0	0.776023	Mn2CrO4_Li_on_Mn_0	14.8787
Mn2CrO4_Cr_on_Mn_0	Mn2CrO4_Cu_on_Mn_0	0.692069	Mn2CrO4_Li_on_Mn_0	14.7162
Mn2CrO4_O_vacancy_0	Mn2CrO4_Cr_on_Mn_0	1.4017	Mn2CrO4_Li_on_Mn_0	14.8178
Mn2CrO4_Li_on_Mn_0	Mn2CrO4_Na_on_Mn_0	8.49961	Mn2CrO4_Li_on_Cr_0	20.4275
Mn2CrO4_Na_on_Mn_0	Mn2CrO4_K_on_Mn_0	1.82446	Mn2CrO4_Li_on_Cr_0	15.8181
Mn2CrO4_Zn_on_Mn_0	Mn2CrO4_Cu_on_Mn_0	1.34267	Mn2CrO4_Li_on_Mn_0	14.778
Mn2CrO4_Rb_on_Mn_0	Mn2CrO4_Mg_on_Mn_0	3.78757	Mn2CrO4_Li_on_Cr_0	15.189
Mn2CrO4_Mg_on_Cr_0	Mn2CrO4_Zn_on_Cr_0	3.43517	Mn2CrO4_Li_on_Mn_0	15.26
Mn2CrO4_Mg_on_Mn_0	Mn2CrO4_Na_on_Mn_0	3.57305	Mn2CrO4_Li_on_Cr_0	14.5951
Mn2CrO4_Cr_vacancy_0	Mn2CrO4_Cu_on_Cr_0	2.97627	Mn2CrO4_Li_on_Cr_0	15.8487
Mn2CrO4_K_on_Cr_0	Mn2CrO4_Na_on_Cr_0	1.95465	Mn2CrO4_Li_on_Mn_0	17.0221
Mn2CrO4_Rb_on_Cr_0	Mn2CrO4_K_on_Cr_0	4.15801	Mn2CrO4_Li_on_Mn_0	16.1827
Mn2CrO4_K_on_Mn_0	Mn2CrO4_Na_on_Mn_0	1.82446	Mn2CrO4_Li_on_Cr_0	16.3614
Mn2CrO4_Cu_on_Cr_0	Mn2CrO4_Ag_on_Cr_0	0.776023	Mn2CrO4_Li_on_Mn_0	14.8241

Table 22: The distances to the nearest and farthest neighbors for the 21 compounds, with charge state 0, obtained from Mn2CrO4.

Compound ID	f-enthalpy	top two nearest neighbors		
		compound	distance	f-enthalpy
Mn2CrO4_Mn_on_Cr_0	1.11166	Mn2CrO4_Cr_on_Mn_0	11.1046	0.921734
		Mn2CrO4_O_vacancy_0	11.1062	5.01357
Mn2CrO4_Ag_on_Mn_0	4.61637	Mn2CrO4_Cu_on_Mn_0	0.701159	3.06201
		Mn2CrO4_Cr_on_Mn_0	1.32404	0.921734
Mn2CrO4_Cu_on_Mn_0	3.06201	Mn2CrO4_Cr_on_Mn_0	0.692069	0.921734
		Mn2CrO4_Ag_on_Mn_0	0.701159	4.61637
Mn2CrO4_Na_on_Cr_0	10.9667	Mn2CrO4_K_on_Cr_0	1.95465	5.07751
		Mn2CrO4_Mg_on_Cr_0	3.51955	-0.242771
Mn2CrO4_Li_on_Cr_0	2.56028	Mn2CrO4_Na_on_Cr_0	8.19348	10.9667
		Mn2CrO4_K_on_Cr_0	8.5689	5.07751
Mn2CrO4_Mn_vacancy_0	5.64584	Mn2CrO4_Cr_on_Mn_0	2.46979	0.921734
		Mn2CrO4_Cu_on_Mn_0	2.67209	3.06201
Mn2CrO4_Zn_on_Cr_0	2.64323	Mn2CrO4_Ag_on_Cr_0	1.93871	6.51794
		Mn2CrO4_Cu_on_Cr_0	1.95502	4.64778
Mn2CrO4_Ag_on_Cr_0	6.51794	Mn2CrO4_Cu_on_Cr_0	0.776023	4.64778
		Mn2CrO4_Zn_on_Cr_0	1.93871	2.64323
Mn2CrO4_Cr_on_Mn_0	0.921734	Mn2CrO4_Cu_on_Mn_0	0.692069	3.06201
		Mn2CrO4_Ag_on_Mn_0	1.32404	4.61637
Mn2CrO4_O_vacancy_0	5.01357	Mn2CrO4_Cr_on_Mn_0	1.4017	0.921734
		Mn2CrO4_Cu_on_Cr_0	1.4785	4.64778
Mn2CrO4_Li_on_Mn_0	0.802801	Mn2CrO4_Na_on_Mn_0	8.49961	9.07765
		Mn2CrO4_K_on_Mn_0	8.76355	3.18072
Mn2CrO4_Na_on_Mn_0	9.07765	Mn2CrO4_K_on_Mn_0	1.82446	3.18072
		Mn2CrO4_Mg_on_Mn_0	3.57305	-2.08582
Mn2CrO4_Zn_on_Mn_0	0.479527	Mn2CrO4_Cu_on_Mn_0	1.34267	3.06201
		Mn2CrO4_Ag_on_Mn_0	1.41205	4.61637
Mn2CrO4_Rb_on_Mn_0	4.32203	Mn2CrO4_Mg_on_Mn_0	3.78757	-2.08582
		Mn2CrO4_K_on_Mn_0	4.17584	3.18072
Mn2CrO4_Mg_on_Cr_0	-0.242771	Mn2CrO4_Zn_on_Cr_0	3.43517	2.64323
		Mn2CrO4_Na_on_Cr_0	3.51955	10.9667
Mn2CrO4_Mg_on_Mn_0	-2.08582	Mn2CrO4_Na_on_Mn_0	3.57305	9.07765
		Mn2CrO4_Zn_on_Mn_0	3.578	0.479527
Mn2CrO4_Cr_vacancy_0	6.65839	Mn2CrO4_Cu_on_Cr_0	2.97627	4.64778
		Mn2CrO4_Cr_on_Mn_0	2.99412	0.921734
Mn2CrO4_K_on_Cr_0	5.07751	Mn2CrO4_Na_on_Cr_0	1.95465	10.9667
		Mn2CrO4_Rb_on_Cr_0	4.15801	6.39256
Mn2CrO4_Rb_on_Cr_0	6.39256	Mn2CrO4_K_on_Cr_0	4.15801	5.07751
		Mn2CrO4_Na_on_Cr_0	4.47093	10.9667
Mn2CrO4_K_on_Mn_0	3.18072	Mn2CrO4_Na_on_Mn_0	1.82446	9.07765
		Mn2CrO4_Rb_on_Mn_0	4.17584	4.32203
Mn2CrO4_Cu_on_Cr_0	4.64778	Mn2CrO4_Ag_on_Cr_0	0.776023	6.51794
		Mn2CrO4_Cr_on_Mn_0	1.46654	0.921734

Table 23: The distances to the nearest two neighbors, and their f-enthalpy values, for the 21 compounds, with charge state 0, obtained from Mn2CrO4.

Compound ID	nearest neighbor		farthest neighbor	
	compound	distance	compound	distance
Mn2CrO4_Mn_on_Cr_-1	Mn2CrO4_Cu_on_Mn_-1	11.1185	Mn2CrO4_Li_on_Mn_-1	18.4639
Mn2CrO4_Ag_on_Mn_-1	Mn2CrO4_Cu_on_Mn_-1	0.701159	Mn2CrO4_Li_on_Mn_-1	15.2968
Mn2CrO4_Cu_on_Mn_-1	Mn2CrO4_Ag_on_Mn_-1	0.701159	Mn2CrO4_Li_on_Mn_-1	14.988
Mn2CrO4_Na_on_Cr_-1	Mn2CrO4_K_on_Cr_-1	1.95465	Mn2CrO4_Li_on_Mn_-1	16.4218
Mn2CrO4_Li_on_Cr_-1	Mn2CrO4_Na_on_Cr_-1	8.19348	Mn2CrO4_Li_on_Mn_-1	20.4275
Mn2CrO4_Mn_vacancy_-1	Mn2CrO4_Cu_on_Mn_-1	2.67209	Mn2CrO4_Li_on_Mn_-1	16.3904
Mn2CrO4_Zn_on_Cr_-1	Mn2CrO4_Ag_on_Cr_-1	1.93871	Mn2CrO4_Li_on_Mn_-1	14.9409
Mn2CrO4_Ag_on_Cr_-1	Mn2CrO4_Cu_on_Cr_-1	0.776023	Mn2CrO4_Li_on_Mn_-1	14.8787
Mn2CrO4_Li_on_Mn_-1	Mn2CrO4_Na_on_Mn_-1	8.49961	Mn2CrO4_Li_on_Cr_-1	20.4275
Mn2CrO4_Na_on_Mn_-1	Mn2CrO4_K_on_Mn_-1	1.82446	Mn2CrO4_Li_on_Cr_-1	15.8181
Mn2CrO4_Rb_on_Mn_-1	Mn2CrO4_K_on_Mn_-1	4.17584	Mn2CrO4_Li_on_Cr_-1	15.189
Mn2CrO4_Mg_on_Cr_-1	Mn2CrO4_Zn_on_Cr_-1	3.43517	Mn2CrO4_Li_on_Mn_-1	15.26
Mn2CrO4_Cr_vacancy_-1	Mn2CrO4_Cu_on_Cr_-1	2.97627	Mn2CrO4_Li_on_Cr_-1	15.8487
Mn2CrO4_K_on_Cr_-1	Mn2CrO4_Na_on_Cr_-1	1.95465	Mn2CrO4_Li_on_Mn_-1	17.0221
Mn2CrO4_Rb_on_Cr_-1	Mn2CrO4_K_on_Cr_-1	4.15801	Mn2CrO4_Li_on_Mn_-1	16.1827
Mn2CrO4_K_on_Mn_-1	Mn2CrO4_Na_on_Mn_-1	1.82446	Mn2CrO4_Li_on_Cr_-1	16.3614
Mn2CrO4_Cu_on_Cr_-1	Mn2CrO4_Ag_on_Cr_-1	0.776023	Mn2CrO4_Li_on_Mn_-1	14.8241

Table 24: The distances to the nearest and farthest neighbors for the 17 compounds, with charge state -1, obtained from Mn2CrO4.

Compound ID	f-enthalpy	top two nearest neighbors		
		compound	distance	f-enthalpy
Mn2CrO4_Mn_on_Cr_-1	2.45537	Mn2CrO4_Cu_on_Mn_-1	11.1185	4.02364
		Mn2CrO4_Ag_on_Mn_-1	11.1646	4.98535
Mn2CrO4_Ag_on_Mn_-1	4.98535	Mn2CrO4_Cu_on_Mn_-1	0.701159	4.02364
		Mn2CrO4_Cu_on_Cr_-1	1.86686	5.09537
Mn2CrO4_Cu_on_Mn_-1	4.02364	Mn2CrO4_Ag_on_Mn_-1	0.701159	4.98535
		Mn2CrO4_Cu_on_Cr_-1	1.56778	5.09537
Mn2CrO4_Na_on_Cr_-1	11.4986	Mn2CrO4_K_on_Cr_-1	1.95465	5.70444
		Mn2CrO4_Mg_on_Cr_-1	3.51955	0.019007
Mn2CrO4_Li_on_Cr_-1	3.11996	Mn2CrO4_Na_on_Cr_-1	8.19348	11.4986
		Mn2CrO4_K_on_Cr_-1	8.5689	5.70444
Mn2CrO4_Mn_vacancy_-1	6.00623	Mn2CrO4_Cu_on_Mn_-1	2.67209	4.02364
		Mn2CrO4_Ag_on_Mn_-1	2.83319	4.98535
Mn2CrO4_Zn_on_Cr_-1	2.94207	Mn2CrO4_Ag_on_Cr_-1	1.93871	7.24689
		Mn2CrO4_Cu_on_Cr_-1	1.95502	5.09537
Mn2CrO4_Ag_on_Cr_-1	7.24689	Mn2CrO4_Cu_on_Cr_-1	0.776023	5.09537
		Mn2CrO4_Zn_on_Cr_-1	1.93871	2.94207
Mn2CrO4_Li_on_Mn_-1	0.995454	Mn2CrO4_Na_on_Mn_-1	8.49961	9.29108
		Mn2CrO4_K_on_Mn_-1	8.76355	3.45242
Mn2CrO4_Na_on_Mn_-1	9.29108	Mn2CrO4_K_on_Mn_-1	1.82446	3.45242
		Mn2CrO4_Rb_on_Mn_-1	4.18869	4.67202
Mn2CrO4_Rb_on_Mn_-1	4.67202	Mn2CrO4_K_on_Mn_-1	4.17584	3.45242
		Mn2CrO4_Na_on_Mn_-1	4.18869	9.29108
Mn2CrO4_Mg_on_Cr_-1	0.019007	Mn2CrO4_Zn_on_Cr_-1	3.43517	2.94207
		Mn2CrO4_Na_on_Cr_-1	3.51955	11.4986
Mn2CrO4_Cr_vacancy_-1	7.27288	Mn2CrO4_Cu_on_Cr_-1	2.97627	5.09537
		Mn2CrO4_Cu_on_Mn_-1	3.04499	4.02364
Mn2CrO4_K_on_Cr_-1	5.70444	Mn2CrO4_Na_on_Cr_-1	1.95465	11.4986
		Mn2CrO4_Rb_on_Cr_-1	4.15801	7.05915
Mn2CrO4_Rb_on_Cr_-1	7.05915	Mn2CrO4_K_on_Cr_-1	4.15801	5.70444
		Mn2CrO4_Na_on_Cr_-1	4.47093	11.4986
Mn2CrO4_K_on_Mn_-1	3.45242	Mn2CrO4_Na_on_Mn_-1	1.82446	9.29108
		Mn2CrO4_Rb_on_Mn_-1	4.17584	4.67202
Mn2CrO4_Cu_on_Cr_-1	5.09537	Mn2CrO4_Ag_on_Cr_-1	0.776023	7.24689
		Mn2CrO4_Cu_on_Mn_-1	1.56778	4.02364

Table 25: The distances to the nearest two neighbors, and their f-enthalpy values, for the 17 compounds, with charge state -1, obtained from Mn2CrO4.

Compound ID	f-enthalpy	nearest neighbor			
		compound	distance	f-enthalpy	difference
Mn2CrO4_Ag_on_Mn_0	4.61637	Mn2CrO4_Cu_on_Mn_0	0.701159	3.06201	1.55436
Mn2CrO4_Ag_on_Mn_-1	4.98535	Mn2CrO4_Cu_on_Mn_-1	0.701159	4.02364	0.96171
Mn2CrO4_Na_on_Cr_0	10.9667	Mn2CrO4_K_on_Cr_0	1.95465	5.07751	5.88919
Mn2CrO4_Na_on_Cr_-1	11.4986	Mn2CrO4_K_on_Cr_-1	1.95465	5.70444	5.79416
Mn2CrO4_Li_on_Cr_0	2.56028	Mn2CrO4_Na_on_Cr_0	8.19348	10.9667	-8.40642
Mn2CrO4_Li_on_Cr_-1	3.11996	Mn2CrO4_Na_on_Cr_-1	8.19348	11.4986	-8.37864
Mn2CrO4_Zn_on_Cr_0	2.64323	Mn2CrO4_Ag_on_Cr_0	1.93871	6.51794	-3.87471
Mn2CrO4_Zn_on_Cr_-1	2.94207	Mn2CrO4_Ag_on_Cr_-1	1.93871	7.24689	-4.30482
Mn2CrO4_Ag_on_Cr_0	6.51794	Mn2CrO4_Cu_on_Cr_0	0.776023	4.64778	1.87016
Mn2CrO4_Ag_on_Cr_-1	7.24689	Mn2CrO4_Cu_on_Cr_-1	0.776023	5.09537	2.15153
Mn2CrO4_Li_on_Mn_0	0.802801	Mn2CrO4_Na_on_Mn_0	8.49961	9.07765	-8.274849
Mn2CrO4_Li_on_Mn_-1	0.995454	Mn2CrO4_Na_on_Mn_-1	8.49961	9.29108	-8.295626
Mn2CrO4_Na_on_Mn_0	9.07765	Mn2CrO4_K_on_Mn_0	1.82446	3.18072	5.89693
Mn2CrO4_Na_on_Mn_-1	9.29108	Mn2CrO4_K_on_Mn_-1	1.82446	3.45242	5.83866
Mn2CrO4_Mg_on_Cr_0	-0.242771	Mn2CrO4_Zn_on_Cr_0	3.43517	2.64323	-2.886001
Mn2CrO4_Mg_on_Cr_-1	0.019007	Mn2CrO4_Zn_on_Cr_-1	3.43517	2.94207	-2.923063
Mn2CrO4_Cr_vacancy_0	6.65839	Mn2CrO4_Cu_on_Cr_0	2.97627	4.64778	2.01061
Mn2CrO4_Cr_vacancy_-1	7.27288	Mn2CrO4_Cu_on_Cr_-1	2.97627	5.09537	2.17751
Mn2CrO4_Rb_on_Cr_0	6.39256	Mn2CrO4_K_on_Cr_0	4.15801	5.07751	1.31505
Mn2CrO4_Rb_on_Cr_-1	7.05915	Mn2CrO4_K_on_Cr_-1	4.15801	5.70444	1.35471

Table 26: The difference between f-enthalpy values for a compound and its nearest neighbor, both with charge states 0, and the corresponding compounds with charge state -1 (which are also nearest neighbors). Note the similarity in the differences, indicating that if three of the values are known, the fourth could be predicted. Compounds obtained from Mn2CrO4.