Ionic Liquid Parameter Prediction Leveraging Quantum Structure Property Relationships

Mitchell Woolever[1] and James Nabity[2]
University of Colorado Boulder, Boulder, CO, 80309

Ronald Cook[3]
MDI LLC, Lakewood, CO, 80403

and

Eric Fox[4]
NASA Marshall Space Flight Center, Huntsville, AL, 35808

U.S. Space Exploration Policy denotes the critical importance of establishing an outpost on the Moon to provide the foundation for human missions beyond cislunar space. However, launching spare components and systems from Earth will likely be cost prohibitive, so the single most important development that is required for enhancing, and in some cases enabling, sustained human presence on the Lunar surface is having the capability to extract metals, oxygen, and water from the Lunar regolith. Ionic Liquids (ILs) are noteworthy for their host of unique chemical properties: a relatively large temperature range in the liquid phase, negligible vapor pressures, thermal and chemical stability, wide voltage window, and many have low toxicity. Furthermore, their coupled organic and ionic nature make them excellent solvents for a wide range of materials. In particular, acidic ionic liquids show the potential to enhance oxygen and metals production from regolith via dissolution and electrolysis. Furthermore, given their organic composition, the physical and chemical properties of ILs can be fine-tuned by modifying their ion structures and combination. Relative abundance changes with sample location, but the principal metals of interest for In Situ Resource Utilization (ISRU) in the Lunar regolith are iron, aluminum, magnesium, calcium, and titanium. However, an IL has yet to be identified that reliably dissolves titanium dioxide or silicon dioxide. Manufacturing and testing even a relatively small subset of the million theoretically stable IL anion/cation combinations for mineral digestion performance analysis is time and cost prohibitive. This paper will discuss a software process pipeline and corresponding analysis setpoints for a method to determine quantum structure property relationships (QSPR), which relate IL molecular structure to chemical function. Using QSPR, hundreds or even thousands of ILs could be assessed for efficacy in regolith ISRU and beyond.

Nomenclature

DFT	=	Density Functional Theory
IL	=	Ionic Liquid
ISRU	=	In Situ Resource Utilization
fLDM	=	fragmented electron Localization Delocalization Matrix
LDM	=	electron Localization Delocalization Matrix
ML	=	Machine Learning

[1] PhD Candidate, Ann and H.J. Smead Aerospace Engineering Sciences, 3775 Discovery Drive, Boulder, CO 80303.
[2] Associate Professor, Ann and H.J. Smead Aerospace Engineering Sciences, 3775 Discovery Drive, Boulder, CO 80303.
[3] Principal Scientist, MDI LLC, 754 S Braun St, Lakewood, CO
[4] Team Lead, EM22 Chemistry & Contamination Control, Martin Rd SW, Huntsville, AL 35808.

MSFC = Marshall Space Flight Center
NIST = National Institute of Standards and Technology
PC = Principal Component
PCA = Principal Component Analysis
QSPR = Quantum Structure Property Relationship
QTAIM = Quantum Theory of Atoms in Molecules
ST = Surface Tension

I. Introduction

THE United States' Space Exploration policy denotes the critical importance of establishing an outpost on the Moon to provide the foundation for human missions beyond cislunar space. The likelihood of contingency scenarios during such a mission necessitates a source of spare parts and tools. However, launching every required spare component and system from Earth is cost prohibitive and fails to account for possible junctures. To reduce component failure down time, recover quickly from equipment operation degradation, improve system flexibility, and ultimately enhance crew safety, a long—term Lunar outpost must have a system to manufacture tools, parts, structures, and spares in situ and on demand[1,2]. Section 7.3.1 of the NASA Space Technology Roadmap specifically states the importance of ISRU processing and production and indicates that extensive work is still required to develop robust systems which produce consumables at rates that meet mission requirements. In particular, acidic Ionic Liquids (ILs) show the potential to enhance oxygen and metals production from regolith and meteoritic metal via dissolution and electrolysis[3]. Furthermore, prior research establishes the feasibility for low temperature iron oxide digestion with an aqueous solution of the acidic IL 1-ethyl-3-methylimidazolium hydrogen sulfate [$Emim^+$][HSO_4^-][4,5,6]. The acidic hydrogen sulfate anion [HSO_4^-] donates protons to facilitate metal oxide digestion. As shown in Equation 1, the IL digests a metal oxide (MO) and drives metal ions into solution. Note that Equation 1 is not balanced for charge or stoichiometry and does not account for metal oxidation state, as the reaction specifics are dependent on the metal oxide of interest.

$$\mathrm{MO_{(s)}} + \mathrm{EMIM^+_{(aq)}} + \mathrm{HSO^-_{4(aq)}} \rightarrow \mathrm{M^+_{(aq)}} + \mathrm{SO^{2-}_{4\,(aq)}} + \mathrm{EMIM^+_{(aq)}} + \mathrm{H_2O} \quad (1)$$

ILs are salts with melting temperatures below 100°C that are generally composed of an organic cation and an inorganic or organic anion[7]. ILs are noteworthy for their host of unique chemical properties, which include a relatively large liquid temperature range and negligible vapor pressures. Furthermore, their coupled organic and ionic nature make them excellent solvents for a wide range of materials[7]. Unlike many other solvents, numerous ILs have low toxicity numbers[5]. Furthermore, given their organic composition, the physical and chemical properties of ILs can be fine-tuned by modifying their cation moieties and cation-anion combination[7].

With at least a million theoretically stable IL ion configurations, creative molecular design opens the way for their application in extraction, membrane infiltration, gas separation, electrolytes, lubricants, and catalysts[7]. However, the range of IL possibilities necessitates development and deployment of task-specific IL property prediction in favor of exhaustive synthesis and characterization. The latter approach would be time and cost prohibitive for even a small subset of all ILs[5].

Relating chemical properties to molecular structure via quantum chemical Quantitative Structure Property Relationships (QSPRs) is a demonstrably successful approach to linking IL structural configurations and chemical properties[8]. This research sought to develop a methodology to accurately predict physical, chemical, and thermodynamic properties for ILs where the data are not yet available. 'How good is good enough' remains to be seen, although correlations with $R^2 \geq 0.8$ are desired. Ultimately, we seek to apply this method as a screening tool to identify ILs that appear most promising for regolith digestion. Initially however, intrinsic IL properties such as surface tension, electrical conductivity, and electrochemical and thermal stability will be targeted.

QSPR enables rapid, spaceflight application-specific IL design and selection. Molecular structure descriptors for modeling and predicting IL properties must be generated to enable QSPR. Machine Learning (ML) algorithms can be employed to establish relationships between an IL's structure and intrinsic properties. ML algorithms perform best when they are trained on large, clean, unbiased datasets, but IL intrinsic property data reported in the academic literature can be sparse, biased toward a few common ions, and in conflict with the values reported by other researchers. A method for efficient IL intrinsic property collection and vetting has been established and is discussed in Section III.

II. Summary of Approach

Density Functional Theory (DFT) is a commonly applied quantum mechanical modeling approach in chemistry, physics, and materials science. As shown in Figure 1, DFT establishes the basis for mapping a chemical's intrinsic thermophysical properties onto its probabilistic electron distribution topology[9]. Descriptors based on the topological analysis of molecular electron density enable QSPR establishment and are the ML feature of emphasis for this modeling effort[8]. A QSPR Modeling Pipeline (Figure 3) has been developed to predict IL thermophysical property information from molecular structure via electron density topology in the form of the electron Localization Delocalization Matrix (LDM), shown in Figure 2. QSPR IL property predictions were validated experimentally at the NASA Marshall Space Flight Center (MSFC) IL Laboratory.

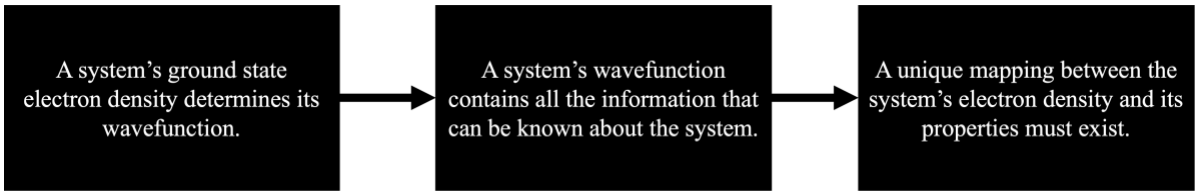


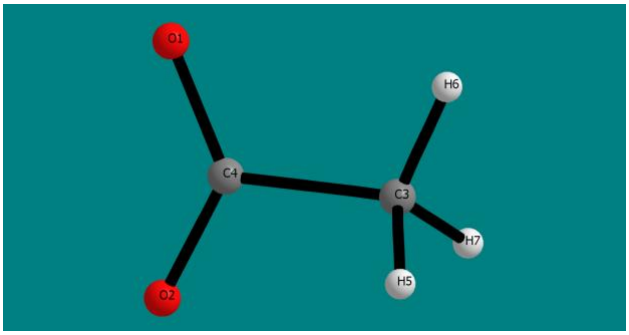
Figure 1. Modeling approach motivation.

III. Modeling Methods

Electron probability density topology can be calculated from an IL's geometry and expressed in an LDM. The LDM enumerates electron sharing between each atom in a molecule and a set of IL LDMs is fed into a ML algorithm to establish a mapping between LDM features and intrinsic IL thermophysical properties. This section expands on the QSPR Modeling Pipeline steps road mapped in Figure 3. The IL QSPR modeling process begins with a molecular structure file, builds a corresponding LDM, and produces thermophysical QSPR insights with a set of training LDM Principal Components (PCs) in conjunction with a learning algorithm.

LDM Calculation: The modeling process begins with a molecular structure file, which contains only data regarding the constituent atomic species and their respective spatial coordinates. The structure files are passed into the Gaussian software, where they are relaxed to their minimum energy geometry. Gaussian also calculates the ground-state wave function using DFT[8]. The ground-state wavefunction is passed into the AIMAll software, which applies the Quantum Theory of Atoms in Molecules (QTAIM) to separate the IL into atom basins. QTAIM treats a molecule as a collection of atoms to which the physics of open systems can be applied. QTAIM's underlying concept is the atom basin, which is defined as a region of space bounded by a surface satisfying the quantum boundary condition of zero-flux in the gradient vector field of electron density[8]. The application of the atom basin boundary condition to a molecular charge distribution yields a complete partitioning of a molecule into non-overlapping atoms, and the calculated properties can be summed to provide global molecular properties. AIMAll calculates the LDM, which is extracted via AIMLDM, a script developed by the Professor Cherif Matta group at Mount Saint Vincent University[10].

Figure 2 illustrates an example LDM and corresponding molecule geometry. The matrix elements enumerate electron sharing between every combination of two atoms in the molecule. The matrix diagonal, which indicates electron 'self-sharing', usually possesses most of the electron density. The sum of the row sums is equal to the total number of electrons in the molecule.



(a) Acetate anion.

	O1	O2	C3	C4	H5	H6	H7	Row Sums
O1	8.459884	0.173636	0.070997	0.56374	0.009234	0.013237	0.009236	9.299964
O2	0.173636	8.466717	0.072542	0.559718	0.010602	0.010754	0.01061	9.30458
C3	0.070997	0.072542	4.006538	0.422854	0.486281	0.485751	0.486274	6.031237
C4	0.56374	0.559718	0.422854	2.667096	0.017666	0.014254	0.017663	4.262991
H5	0.009234	0.010602	0.486281	0.017666	0.466985	0.020932	0.021999	1.033699
H6	0.013237	0.010754	0.485751	0.014254	0.020932	0.46793	0.020931	1.033788
H7	0.009236	0.01061	0.486274	0.017663	0.021999	0.020931	0.466977	1.033691
							Electron #	31.99995

→ = sum operation

(b) Acetate LDM and electron sum.

Figure 2. The ball-and stick model for the acetate anion (a) and corresponding LDM (b).

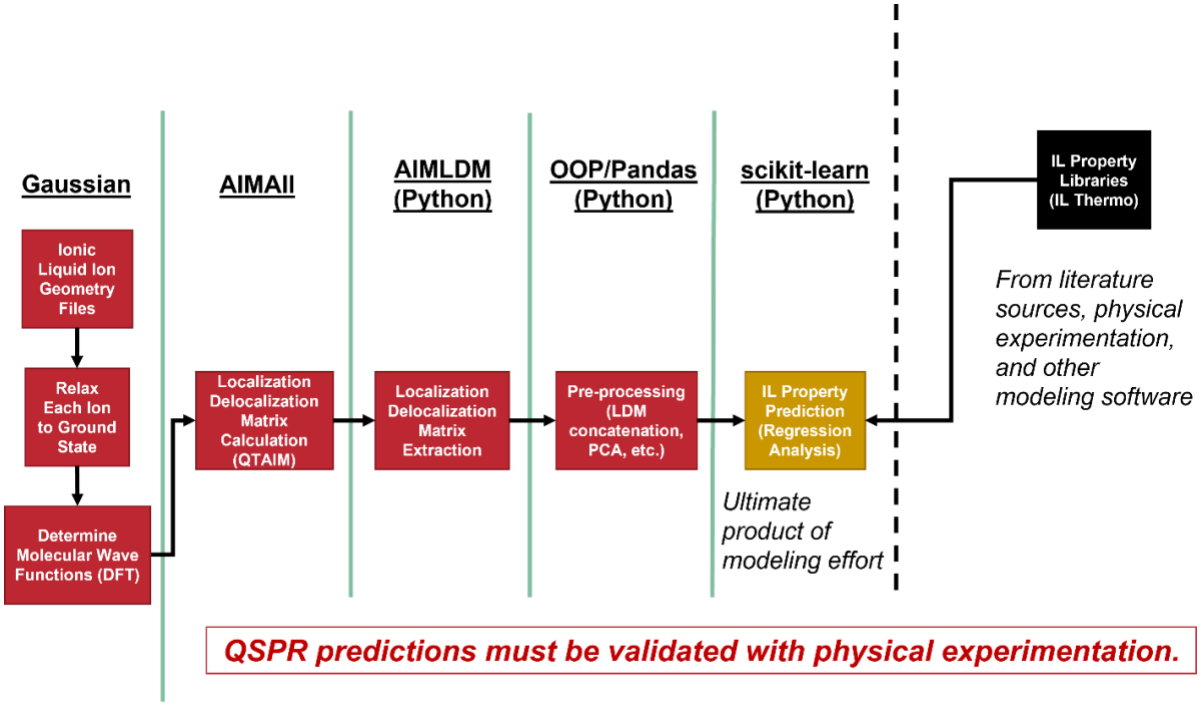


Figure 3. IL QSPR modeling pipeline.

Dimensional Reduction: Most machine learning regression algorithms do not perform well with high-dimensional data. The number of variables in an LDM is equal to the triangular number corresponding to the atom count in the molecule. Large IL ions contain dozens of atoms, which results in an untenable number of LDM variables. Furthermore, regression algorithms must compare data of equal dimensionality. It is not clear how to slice or truncate larger LDMs and retain representative integrity such that their electron topology is enumerated with an equal number of variables as smaller LDMs in the same data set. Smaller LDMs could be buffered with gaussian noise or null values to match the largest LDM in a data set, but the issue of high dimensionality remains.

Principal Component Analysis (PCA) is a method of dimensional reduction which projects a set of orthogonal vectors equal in number to the dimension of a matrix onto the matrix data by maximizing the sum of squared distances. This process establishes the matrix eigenspace. The eigenvectors are the matrix PCs. By finding the sum of square distances for each eigenvector, the set of eigenvalues can be calculated. These eigenvalues' corresponding magnitudes indicate the comparative data variance to be attributed to each PC[11]. The number of principal components for an LDM is equal to the number of atoms in the corresponding molecule. Larger molecules' LDMs can be truncated to match the vector length of the smallest IL in a dataset without significant penalty, as each subsequent principal component accounts for increasingly less variance in the electron distribution probability topology. By conducting PCA on the covariance matrices of IL LDMs, the shape and charge density of an IL's electron cloud can be completely and succinctly quantified. Machine learning algorithms can correlate LDM principal component eigenvalues with some thermophysical properties[12].

LDM Fragments: The application of fragmented LDMs (fLDMs), an approach that leverages the superposition principle at play in electron probability density, significantly reduces the Gaussian software's ground state molecular geometry calculation time. fLDM employment is vital as wavefunction determination for large or complex molecules can take days. Figure 4 illustrates the concept of fLDM construction. fLDMs are constructed by calculating the corresponding ground-state wavefunctions, atom basins, and LDMs for an IL's ions separately (as fragments). The LDM fragments are concatenated diagonally, and blank spaces are buffered with null values.

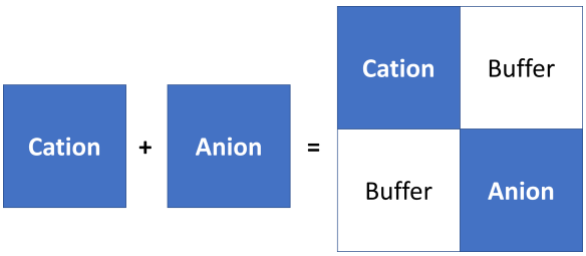


Figure 4. fLDM concatenation for an IL.

A comparison in ground state wavefunction calculation times for complete ILs versus IL fragments is shown in Table 1. Calculating the ground state wavefunctions for IL ions separately averages to a 12-fold decrease in calculation time with an average PC difference between the LDM and fLDM of 0.33% in this set of six ILs. Very little data is lost, and weeks of calculation time is saved through the application of fLDMs when working with hundreds of ILs.

Table 1. LDM vs. fLDM calculation time and principal component comparison

Ionic Liquid	LDM Time (hr)	fLDM Time (hr)	Execution Time Ratio	# of e^-	PC R^2	% PC difference
1-Ethyl-3-methylimidazolium Hydrogen sulfate	9.107	0.7583	**12.0**	110	0.9998	**0.012**
1-Methylimidazolium Hydrogen sulfate	3.603	0.1175	**30.7**	94	0.9998	**0.014**
Glutamic Bistrifluoromethanesulfonyl)imide	59.72	6.13	**9.74**	216	0.9999	**0.008**
Glutamic Hydrogen sulfate	2.933	2.706	**1.08**	128	0.9993	**0.063**
Taurinium Hydrogen sulfate	1.534	0.555	**2.76**	116	0.9786	**2.14**
Trimethyl glycinium Hydrogen sulfate	20.39	1.176	**17.3**	114	0.9999	**0.01**

Machine Learning: Quantum theory establishes the correlation between a molecule's electron probability density topology and its intrinsic properties. However, these relationships are opaque and likely highly nonlinear. ML regression is applied to automate the establishment of QSPRs between LDM PCs and IL intrinsic properties. An ML process likely to provide users with valuable and desirable insights follows the design sequence detailed in Table 2[13].

IL Properties and Collection: For their employment in regolith ISRU, the primary intrinsic IL properties of interest for QSPR Modeling Pipeline prediction include the Hansen Solubility Parameters and other solvation-related properties, like surface tension. However, transport properties, like viscosity and electrical conductivity, and chemical stability characteristics, like thermal and electrochemical window, have systems-level implications that will prove useful for parameterizing an applied ISRU process.

The National Institute of Standards and Technology (NIST) maintains the most detailed, comprehensive, and navigable IL property database to date: IL Thermo[14]. Despite this repository's utility, manual data collection of a

Table 2. A philosophy for deriving insight from ML frameworks.

ML Development Step	Explanation
1. Exploratory Analysis	Initial familiarization with data qualities like variance, central tendency, data type, potential biases, data quality, and simple trends. E.g., academic IL property data is disproportionately abundant for imidazolium-based cations and the TFSI anion.
2. Data Cleaning	Outliers and unreliably sourced data are removed. Data is partitioned. E.g., the QSPR pipeline uses outlier detection methods like isolation forests.
3. Feature Engineering	Determine which data features are most correlated with the desired insights. E.g., the QSPR pipeline uses LDM PCs as an expression of an IL's wavefunction.
4. Algorithm Selection	Selection between supervised and unsupervised learning and between classification and regression. A wide array of algorithms is available in each of these categories. Each is suited to different data set sizes, distributions, and qualities. E.g., the QSPR pipeline uses a host of competing, supervised regression algorithms in parallel.
5. Model Training	Establish a test / train partition and tune algorithm hyperparameters. Iterate regression. E.g., the QSPR pipeline uses an 80/20 train/test partition, so regression algorithms learn off 80% of the data and are tested against the remaining 20%.

single property for a few hundred ILs is painstaking and can take weeks. A web scraper to automate IL property collection from IL Thermo has been developed. Physical and thermophysical data for hundreds of ILs are now available for QSPR investigation on demand. In this investigation, properties from IL Thermo are regressed against a set of IL fLDM PC eigenvalues via a machine learning algorithm. To date, the best QSPR Modeling Pipeline fits have been produced while predicting IL surface tension.

IV. Experimental Methods

Models are only as good as their pragmatic utility. While a portion of a training dataset can be isolated from a ML algorithm to test model efficacy on novel data, the purpose of this investigation is to develop the capability to predict thermophysical properties of uncharacterized ILs. This capability was demonstrated by measuring Surface Tension (ST) for a set of ILs (some of which did not have a ST value reported in the literature) and comparing the predicted value from the modeling pipeline with the experimentally obtained value. This work was conducted in the MSFC IL Laboratory.

ST data was taken for a total of 10 ILs via the Du Nüoy ring method (Surface Tensiometer Model 20, Fisher Scientific). These included: 1-ethyl-3-methylimidazolium dicyanamide (>98% purity, Iolitec), 1-butyl-3-methylimidazolium acetate (>98% purity, Iolitec), 1-butyl-1-methylpyrrolidinium triflate (99% purity, Iolitec), 1-ethyl-3-methylimidazolium hydrogen sulfate (95% purity, Sigma Aldrich), ethylammonium nitrate (synthesized by IL Sciences Group at MSFC), 1-butyl-3-methylimidazolium tetrafluoroborate (Oakwood Chemical), 3-(butyl-4-sulfonic acid)-1-methylimidazolium hydrogen sulfate (synthesized by IL Sciences Group at MSFC), tributyl(dodecyl)phosphonium benzotriazolide (synthesized by IL Sciences Group at MSFC), tributyl(tetradecyl)phosphonium 3-(trifluoromethyl)pyrazolide (synthesized by IL Sciences Group at MSFC), 3-(butyl-4-sulfonic acid)-1-methylimidazolium triflate (synthesized by IL Sciences Group at MSFC).

The ST measurement setup is pictured in Figure 5. Before each IL ST measurement, the platinum Du Nüoy ring was rinsed with isopropanol, benzene, and acetone. The remaining surface organics were then driven off with an acetylene torch. The ring was rinsed in 17.5 MΩ deionized water. IL sample watch glasses were etched in an aqueous solution of NaOH and isopropanol for ~45 min prior to sample filling. ILs underwent thermal vacuum bake-out for at 24 - 48 hr prior to ST measurement at ~50°C and ~31 mTorr. Additional IL purification steps such as activated charcoal filtering were not taken. The thermal vacuum oven was backfilled with dry Ar gas during repressurization, and individual IL samples were not retrieved until directly prior to ST measurement. IL water content was assessed directly after retrieval from the vacuum oven and directly after ST measurement via Karl Fischer titration (V20S Volumetric KF Titrator, Mettler Toldedo). IL temperature was controlled with a chiller (RTE-111, Neslab) and measured during ST data collection via a submerged thermocouple probe. Ambient air temperature and RH were measured (AcuRite) for two-phase tensiometer buoyancy correction.

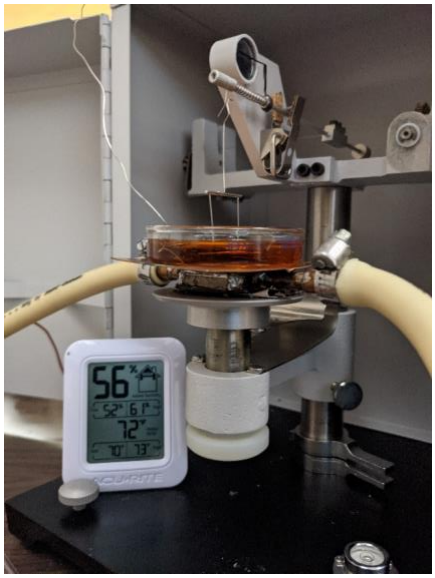


Figure 5. Temperature-controlled IL surface tension measurement setup.

V. Results

The results from the IL ST data collection effort are listed in Table 3. Prior to experimental ST collection, the QSPR Modeling Pipeline produced a predictive model for surface tension using 43 ILs (80-20 train-to- test ratio, 37 after outlier cleaning) with a determination coefficient of 0.96 and a 5-k-fold cross validation of 0.55 using a random forest regression algorithm. Data used to train the model was collected from IL Thermo[14]. This is the best property prediction performance the pipeline has demonstrated to date. QSPR Modeling Pipeline predictions for the experimental IL set agreed with the model's predicted values with an R^2 of 0.8, which satisfies the notional method utility requirement. These experimental validation results are illustrated in Figure 6. 1-Butyl-3-methylimidazolium acetate was excluded from the analysis due to its high pre and post-data-collection water content. Experimental

Table 3. Surface Tension measurements and pre / post-measurement IL water content.

Ionic Liquid	Surface Tension (dynes/cm)		Water Content (% mass)		No. of Measurements (ST)
	Average	*Standard Dev.*	*Before*	*After*	
1-Ethyl-3-methylimidazolium Dicyanamide	**55.62**	**1.26E+00**	**0.2762**	**1.2293**	15
	25.7°C	6.96E-02			
1-Butyl-3-methylimidazolium Acetate	**38.16**	**2.45E-01**	**2.4323**	**4.1234**	10
	25.8°C	6.71E-02			
1-Butyl-1-methylpyrrolidinium Triflate	**35.72**	**7.48E-02**	**0.0502**	**0.5392**	10
	25.4°C	6.40E-02			
1-Ethyl-3-methylimidazolium Hydrogen sulfate	**52.07**	**1.79E-01**	**0.5117**	**0.5942**	10
	25.4°C	0.00E+0			
Ethylammonium Nitrate	**47.77**	**7.81E-02**	**0.1676**	**0.3233**	10
	25.5°C	3.07E-01			
1-Butyl-3-methylimidazolium Tetrafluoroborate	**44.26**	**7.00E-02**	**0.0824**	**0.1791**	10
	25.2°C	6.32E-02			
3-(Butyl-4-sulfonic acid)-1-methylimidazolium Hydrogen sulfate	**38.41**	**3.32E-01**	**0.4769**	**0.7975**	10
	25.8°C	4.58E-02			
Tributyl(dodecyl)phosphonium Benzotriazolide	**33.03**	**1.00E-01**	**0.1592**	**1.0918**	10
	24.9°C	7.48E-02			
Tributyl(tetradecyl)phosphonium 3-(Trifluoromethyl)pyrazolide	**30.50**	**6.71E-02**	**0.1881**	**0.3955**	10
	24.7°C	7.00E-2			
3-(Butyl-4-sulfonic acid)-1-methylimidazolium Triflate	**45.96**	**2.83E-01**	**0.2618**	**0.3067**	5
	25.42°C	4.00E-02			

validation results indicate that the developed QSPR modeling pipeline shows promise in providing valuable insights regarding novel IL intrinsic thermophysical properties.

As a machine learning algorithm is trained on more data, its insights tend to improve. However, despite the promising surface tension prediction result, the model's predictive capabilities significantly decreased when the IL Thermo web scraper was applied to expand the surface tension learning set from about 40 to about 250. There are a few possible explanations for this phenomenon. IL thermophysical properties are highly sensitive to contaminants such as water or CO_2 and synthesis purity is not guaranteed. It is possible that some of the data reported in IL Thermo was collected with unsound IL handling and purification methods or that the ILs studied were contaminated with synthesis reactants. An issue like this would be difficult to address, short of procuring and conscientiously characterizing a broad and diverse IL set in-house. Problems in the application of the larger data set could mean that PC analysis is an unsound dimensional reduction method for comparison across species due to the fact that eigenvectors for different IL LDM covariance matrices will establish dissimilar bases from which the regressed eigenvalues are calculated. However, if it is indeed invalid to compare IL LDM PCs across species, the PC regression method should not have worked at all. It is possible that the high predictive capability of the current QSPR Modeling Pipeline for the small surface tension dataset is due to LDM basis

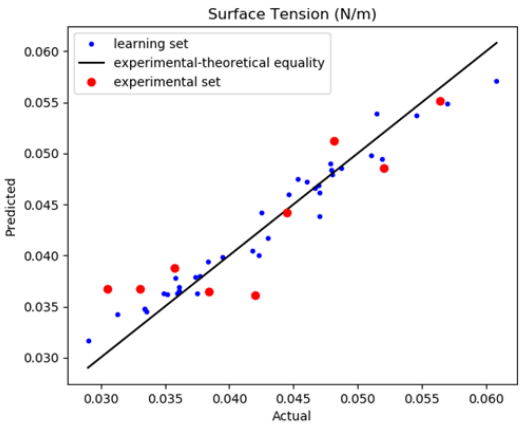


Figure 6. Using fLDM PCs, the QSPR Modeling Pipeline was able to predict the MSFC surface tension experiment results with a determination coefficient of 0.8.

similarities resulting from similar electron cloud topologies in imidazolium ILs (which share a cation structure based on an imidazolium ring), of which the best-fit regression data set mostly comprises. Of the 37 ILs used in the ML training set, 33 cations are imidazolium-based, 3 are pyridinium-based, and 1 is pyrrolidinium-based. The training anions comprise 16 unique species with little structural interrelation. On the other hand, the experimental IL set was selected for structural diversity, and QSPR experimental property prediction remains high.

Further work is required prior to the incorporation of complete IL Thermo data sets for property prediction. Future efforts will explore steps to improve QSPR correlations for larger and more diverse IL data sets by way of IL Thermo data vetting and ML process refinement. IL electron probability distribution topology features other than LDM covariance matrix PCs will be assessed for value as ML features. For instance, property classification with convolutional neural networks on complete IL LDMs may perform more favorably on large, dissimilar IL datasets than direct regression of LDM PCs. Additionally, learning sets could be segregated by base cation to compare against the predictive results produced by a bulk, unsegregated dataset. Furthermore, it is recommended that steps beyond thermal vacuum bake-out be taken to ensure IL purity prior to property measurement, such as activated charcoal filtration.

VI. Conclusion

An experimentally validated modeling pipeline for IL property prediction has been developed and approach feasibility has been established. However, it is unclear whether the approach's predictive capability is sufficient for rapid design and selection of task specific ILs based on Lunar regolith ISRU IL digestion efficacy. Future efforts will explore steps to improve QSPR correlations for larger and more diverse IL data sets by way of IL Thermo data vetting and pre-processing along with ML pipeline refinement. The improved QSPR modeling pipeline will be leveraged to predict other IL properties relevant to regolith ISRU.

Acknowledgments

This work was supported by funding from the NASA NSTRF/NSTGRO Program. The authors would like to thank David Donovan, Dr. Joseph Fillion, Dr. Matthew Marone, Dr. William Kaukler, Christopher Henry, Taylor Middlebrooks, Dr. Nagi Gebraeel, and Dr. Cherif Matta for their assistance throughout this research effort.

References

[1]C. A. McLemore, J. C. Fikes, K. S. McCarley, J. E. Good, J. P. Kennedy, and S. D. Gilley, "From Lunar Regolith to Fabricated Parts: Technology Developments and the Utilization of Moon Dirt," in *Earth & Space 2008*, Long Beach, California, United States: American Society of Civil Engineers, Sep. 2008, pp. 1–1. doi: 10.1061/40988(323)132.

[2]E. Edmunson and C. A. McLemore, "IN SITU MANUFACTURING IS A NECESSARY PART OF ANY PLANETARY ARCHITECTURE," p. 2, 2012.

[3]"NASA Technology Roadmaps - TA 7: Human Exploration Destination Systems," p. 224, 2015.

[4]M. Marone, M. S. Paley, D. N. Donovan, and L. J. Karr, "Lunar Oxygen Production and Metals Extraction Using Ionic Liquids," p. 28, 2009.

[5]E. A. Barrios, P. A. Curreri, and L. J. Karr, "Oxygen Extraction from Regolith Using Ionic Liquids," p. 6, 2011.

[6]L. Karr *et al.*, "Ionic Liquid Facilitated Recovery of Metals and Oxygen from Regolith," in *2018 AIAA SPACE and Astronautics Forum and Exposition*, Orlando, FL: American Institute of Aeronautics and Astronautics, Sep. 2018. doi: 10.2514/6.2018-5291.

[7]P. Nancarrow and H. Mohammed, "Ionic Liquids in Space Technology - Current and Future Trends," *ChemBioEng Rev.*, vol. 4, no. 2, pp. 106–119, Apr. 2017, doi: 10.1002/cben.201600021.

[8]R. Cook, J. A. Nabity, and J. W. Daily, "Characterizing Propellants for Variable-Thrust/Specific Impulse Colloid Thrusters," *J. Propuls. Power*, vol. 33, no. 6, pp. 1325–1331, Nov. 2017, doi: 10.2514/1.B36495.

[9]C. F. Matta and A. A. Arabi, "Electron-density descriptors as predictors in quantitative structure–activity/property relationships and drug design," *Future Med. Chem.*, vol. 3, no. 8, pp. 969–994, Jun. 2011, doi: 10.4155/fmc.11.65.

[10]I. Sumar, R. Cook, P. W. Ayers, and C. F. Matta, "AIMLDM: A program to generate and analyze electron localization-delocalization matrices (LDMs)," 2015.

[11]J. Starmer, "Principal Component Analysis (PCA), Step-by-Step," *StatQuest*, [Online]. Available: https://www.youtube.com/watch?v=FgakZw6K1QQ&t=950s

[12]R. L. Cook, “Principal components of localization-delocalization matrices: new descriptors for modeling biological activities of organic compounds. Part I: mosquito insecticides and repellents,” *Struct. Chem.*, vol. 28, 2017.
[13]EliteDataScience, “Data Science Primer,” 2020, [Online]. Available: https://elitedatascience.com/
[14]NIST, “Ionic Liquids Database - ILThermo.” https://ilthermo.boulder.nist.gov/ (accessed Sep. 08, 2019).