A Novel Modeling Framework for Identifying Environmentally Safe Solvents for Polymer Recycling

Grant Proposal

Jasmine Palit

Massachusetts Academy of Math and Science at the Worcester Polytechnic Institute

Worcester, MA

Author Note

Thank you to Dr. C for supporting me throughout this project and providing me with ample feedback on my writing. Thank you to Dr. Timko and Muntasir Shahabuddin for guiding me in the process of this project and allotting me space to conduct my research.

Executive Summary

Plastic creates large tolls on the environment and even human health through its production and end-of-life effects. To combat this, recycling is used to repurpose old plastics into new products. However, some common plastics, such as multilayer films, have multiple types of polymers in a distinct layered structure. Thus, to recycle them, the layers must first be separated before they can be reused, one method of which is dissolution and precipitation.

During dissolution and precipitation, one polymer is dissolved while the others remain solid, separating them. While this technique is effective, the solvents used are often harmful to the environment and the workers, such as using Toluene to dissolve polyethylene (PE) and polypropylene (PP). Thus, there is a need for more environmentally safe solvents in recycling processes.

This project aims to find safer solvents for PE and PP, given that they presently do not have a safe alternative and are quite common. To accomplish this, this project will utilize Hansen Solubility Parameters as a method deciding which solvents PE and PP will be miscible in. These parameters quantify the idea of 'like dissolves like' by providing a value for each of the main intermolecular forces: dispersion forces, dipole-dipole forces, and hydrogen bonding. However, these parameters are not defined for all green solvents. Thus, this project will use machine learning to predict these parameters for a wide variety of green solvents. Then, the parameters can be compared to the target polymer to determine which environmentally safe solvent is best for dissolving the polymer. After finding the correct solvent, its efficacy can be tested and evaluated.

The implications of this project are creating a healthier planet and reducing human health risks from working with toxic solvents. This project also helps minimize the amount of plastic that is discarded and needs to be produced. The process itself can eventually be applied to any polymer that is in need of a safe solvent, creating better and safer recycling processes overall.

Keywords: Polymer Recycling; Dissolution and Precipitation; Hansen Solubility Parameters

A Novel Modeling Framework for Identifying Environmentally Safe Solvents for Polymer Recycling

Plastic has become an integral part of daily life, due to its wide variety of properties and applications. However, it was estimated in 2015 that 55% of plastic was discarded, leading to adverse environmental effects (Jiang et al., 2020). Plastics do not break down well in the environment by microorganisms like other organic compounds. In turn, this leads to pollution and the creation of nanoplastics which can have negative impacts on ecosystems and human health. Plastic waste also accounts for around 2% of global CO2 emissions, furthering environmental costs (Vollmer et al., 2020). Thus, plastic waste in landfills and the environment must be reduced to protect the environmental health of the planet.

Recycling

Recycling is often used to combat the plastic that has long been piling up in our environment. It was estimated in 2015 that only 20% of plastic waste was recycled (Jiang et al., 2020) ---the rest either left as waste or incinerated. A wide variety of methods exist to recycle plastic, and more are emerging as this problem gets more pressing (Vollmer et al., 2020). However, progress needs to be made to ensure that more waste goes through these processes. Not only does non-recycled plastic pollute the environment but not reusing plastic means needing to produce more. Each year, an estimated 400 million tons of greenhouse gases are released into the atmosphere through the production of new plastics (Shershneva, 2021). Thus, progress needs to be made in the field of recycling to better combat these issues.

Mixed Plastics

One issue with current recycling processes is that they do not always have the capacity to separate polymers. Separation is often necessary for the reproduction of new materials that have specific polymer components. A specific example of multi-polymer plastic is the multilayer film, of which 40% become post-industrial waste after manufacturing (Walker et al., 2020). Multilayer films are films made of stacked layers of polymers. Their layered structure allows a material to benefit from specific properties from each layer, such as high barriers to gases (Tartakowski, 2010). Over 100 million tons of multilayer thermoplastics are produced each year around the world (Walker et al., 2020) making up around 17% of global plastic production. It is highly likely that most of these films do not get recycled due to necessary complexities in the recycling process for multilayer films (Adam et al., 2025). Researchers have been forced to look deeper into how to recycle these films. A promising method to accomplish multilayer film recycling involves selective dissolution and precipitation of individual polymers with targeted solvents.

Dissolution and Recovery of Polymers

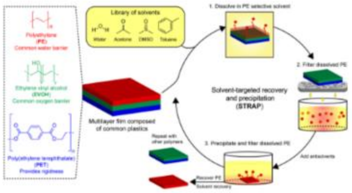


Figure 1. STRAP process of dissolution and precipitation (Walker et al., 2020).

Dissolution for the separation of polymers is a process that is effective at separating chemically different compounds by tailoring solvents to specifically dissolve only one of the compounds present, leaving the others solid. A specific example of this process is the recycling of multilayer films, which are films made of stacked layers of polymers. However, they must be separated into their original polymer components to be reused. Selective dissolution/precipitation has been proven as a successful method of separating the layers, as shown in the method called Solvent Targeted Recovery and Precipitation (STRAP). This process uses a series of solvents compatible with a single layer of the film to dissolve each layer of the film separately and then recover each polymer from its respective solvent (Walker et al., 2020). However, the process described in this paper does not dive deep into the lingering impacts of using toxic solvents to chemically recycle plastic.

Toxic Solvents

One of the primary issues associated with using the process of dissolution is the toxicity of the solvents. Commonly, solvents can cause harm due to toxicity in a variety of diverse ways. They can accidentally or intentionally be released into the environment, risk the safety of those working with them, and potentially remain in the plastics treated with them which go on to be used in industries such as food packaging (Winterton, 2021). The STRAP process defined by Walker et al. (2020) as mentioned above utilized toluene to dissolve polyethylene (PE). However, this solvent poses serious health hazards such as skin irritation, mutagenicity, and reproductive toxicity (Ikegwu et al., 2025; “Toluene - overview”, 2015). Thus, there is a call for more environmentally friendly solvents to be used in recycling processes, especially with PE and PP. Thermodynamic capabilities used in these studies also highlight a pathway that can be used to find suitable solvents for specific polymers, such as the use of Hansen Solubility Parameters (HSPs) (Walker et al., 2020). Using these, nontoxic solvents can be deduced for PE and PP.

Thermodynamic Predictors

Thermodynamic predictors look at molecular properties to predict solubility. Solubility is mostly determined by inter-particle forces and the attraction or repulsion of electrons. This phenomenon can be quantified in a variety of ways. One system, Conductor Like Screening Model for Real Solvents (COSMO-RS) uses the surface charge distributions of molecules and simulates them in different solvents to predict solubility (Klamt, 2011). While this system is accurate, modeling each polymer in every subsequent solvent is time consuming, as the program must simulate the electron charge distribution. An alternative option is the Hansen Solubility Parameters (HSPs). These parameters quantify values for each of the main inter-particular forces — dispersion forces, dipole-dipole forces, and hydrogen bonding — that can then be plotted in 3D space with each axis as a parameter (Abbott, 2025). Based on the general principle of ‘like dissolves like,’ solvents with similar coordinates to polymers are generally miscible. Thus, solvents with similar parameters to polymers can generally dissolve them as they have similar properties. This approach is much more simplistic as solubility does not have to be calculated with every combination of polymer and solvent; rather, the two can be directly compared by their numerical values (Abbott, 2025).

Other parameters that prove useful in predicting the solubility of certain substances are the Abraham accepting and donating parameters. These parameters represent a substance’s ability to donate hydrogens or accept them in a hydrogen bond. However, these parameters have only been defined for a small number of solvents and polymers, calling for a method to predict their values for other systems.

Computer Modeling

Computational programs have been increasingly common in chemistry and predict molecular properties. In the past, others have tried to predict HSPs (Al-Sakkari et al., 2025). However, this research only looked at a small number of parameters for possible training (18 provided by the RDKit software). Conversely, programs like the Mordred molecular descriptor calculator (Moriwaki et al., 2018) offer over 1800 molecular descriptors, creating a much more comprehensive view of a molecule. Thus, deeper exploration into the prediction of HSPs is beneficial for selection of solvents for recycling processes.

Section II: Specific Aims

The overall aim of this project is to create a method to devise what environmentally friendly solvents will be effective at dissolving certain polymers by incorporating machine learning on polymer thermodynamics and validating with physical testing. Our long-term goal is to replace current toxic solvents with more environmentally friendly and less hazardous ones where the central hypothesis is that thermodynamic parameters such as Hansen Solubility Parameters will be able to accurately predict solvents in which these polymers are soluble. The rationale is that using machine learning to predict parameters in certain solvents will reduce the time, cost, labor, and tediousness of testing endless combinations of green solvents on different polymers. This project will also help to create an understanding of which chemical properties most affect HSPs and solubility overall. The work we propose here will create a workflow which can be used to dissolve PE and PP, as well as other polymers in the future. In turn, this process can help reduce environmental harm by minimizing unused plastic waste, preventing the need for more virgin plastics to be produced, and limiting the use of toxic solvents in recycling processes.

Specific Aim 1: The first aim of this project is to collect data on chemicals with predefined HSP values. This step includes gathering data from databases as well as using project-specific APIs.

Specific Aim 2: The second aim of this project is to determine which variables most affect each other and the target HSP. This step will downsize the data needed to calculate each HSP.

Specific Aim 3: The third aim of this project is to develop a machine learning algorithm to predict the HSPs of a variety of different green solvents and their target polymers. Using the refined dataset, a variety of models are tested to determine which approach is most accurate.

Specific Aim 4: The final aim of this project is to test the model results with actual solvents. The target polymers will be dissolved by the solvents predicted by the model. Then, analysis will be conducted on the experimental versus expected solubility.

Expected Outcome: The expected outcome of the project is a workflow where users can enter their polymer. Then, the program will find all necessary data on the polymer and use the models to predict its HSPs. Finally, the model will compare the HSPs of the polymer to potential solvents and solvent belnds and output the best possible solvent mixtures.

Section III: Project Goals and Methodology

Relevance/Significance

Firstly, this project will demonstrate an ability to use non-toxic solvents in recycling processes, thereby reducing the associated environmental costs. Secondly, it will provide a framework for other solvents to be replaced with environmentally safe solvents. Thirdly, by reducing the use of toxic solvents, workers no longer be as exposed to toxic substances, diminishing human health hazards. Finally, by improving recycling processes, this project will allow more plastic to be recycled, thereby minimizing the amount of virgin plastic that must be produced, reducing the energy and greenhouse gas costs used in those processes.

Innovation

This project, unlike past research, focuses on HSPs as the main method of predicting solubility for a more parsimonious approach. This method also uses a larger range of parameters to describe the solvents for training as well as focusing exclusively on green solvents.

Methodology

The project consists of four major steps: collecting data, analyzing data, making predictions, and validating results. The first step is completed by collecting data on the 1000 or so solvents with predefined HSPs as well as a list of green solvents. This data was collected from NistChem, PubMed, and molecular descriptors were sourced through the Mordred Molecular Calculator (Moriwaki et al., 2018). Once collected, the data is analyzed and down –selected to the most important features to predict the HSPs. This process involves using multiple correlation coefficients as well classifying features as either true or false based on their usefulness in prediction of the target variable, completed by BorutaPy. After the features have been chosen to predict each HSP, an variety of models a trained to find the most accurate on, specific to each HSP. This step can be done efficiently using an API called LazyPredict which looks at over 40 models and ranks their accuracies. Finally, models are trained to predict each HSP, and data is collected on their training accuracy. These models are then refined to improve accuracy. To validate the model's results, the predicted solvents are tested on the actual polymers to see how well they dissolve the target polymer. Thus, these steps create a comprehensive methodology to determine the best environmentally safe solvents for the target polymers PE and PP.

Specific Aim #1:

The first aim of this project is to collect and analyze data for a variety of toxic and non-toxic solvents including the one thousand or so solvents with predetermined Hansen Solubility Parameters (HSPs). This data will include information such as SMILES strings, which are strings of information that represent the pieces of the chemical including atomic structure and bonding. These strings can then be fed into the Mordred molecular calculator, which returns over 1800 chemical properties. The rationale for these steps is that they will provide measures for which a model can train on to predict HSPs for chemicals they are currently undefined for.

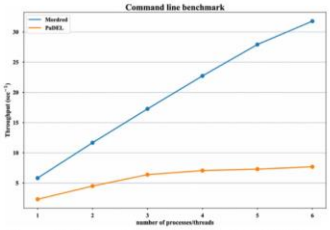
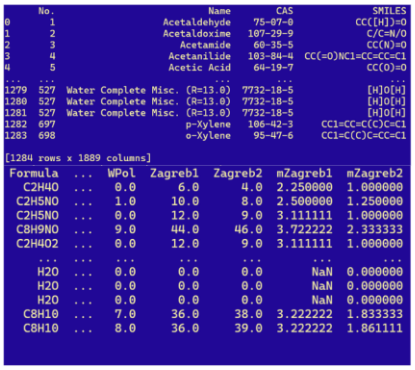


Figure 2. Comparison of Mordred to PaDEL on throughput for 1 to 6 threads. Mordred has a higher throughput overall, and the difference becomes exacerbated as the number of threads increases.

Justification and Feasibility: The methods of this aim are relevant to the overall goal of collecting a wide variety of data to train on. Thus, Mordred provides an exhaustive list of molecular structure description; the program returns over 1800 molecular descriptors. Figure 1 displays how the Mordred calculator compares to one of its competitors, PaDEL. Mordred scored higher on the command line index test, proving that the program is more efficient than its competitor. Overall, the application runs about twice as fast as PaDEL. Thus, Mordred proved to be the best choice for this specific aim.



```
      No.                         Name       CAS               SMILES
0       1                  Acetaldehyde   75-07-0            CC([H])=O
1       2                  Acetaldoxime  107-29-9              C/C=N/O
2       3                     Acetamide   60-35-5              CC(N)=O
3       4                   Acetanilide  103-84-4  CC(=O)NC1=CC=CC=C1
4       5                   Acetic Acid   64-19-7              CC(O)=O
...   ...                           ...       ...                  ...
1279  527  Water Complete Misc. (R=13.0)  7732-18-5             [H]O[H]
1280  527  Water Complete Misc. (R=13.0)  7732-18-5             [H]O[H]
1281  527  Water Complete Misc. (R=13.0)  7732-18-5             [H]O[H]
1282  697                     p-Xylene   106-42-3    CC1=CC=C(C)C=C1
1283  698                     o-Xylene    95-47-6    CC1=C(C)C=CC=C1

[1284 rows x 1889 columns]
Formula  ...  WPol  Zagreb1  Zagreb2  mZagreb1  mZagreb2
   C2H4O  ...   0.0      6.0      4.0  2.250000  1.000000
  C2H5NO  ...   1.0     10.0      8.0  2.500000  1.250000
  C2H5NO  ...   0.0     12.0      9.0  3.111111  1.000000
  C8H9NO  ...   9.0     44.0     46.0  3.722222  2.333333
  C2H4O2  ...   0.0     12.0      9.0  3.111111  1.000000
     ...  ...   ...      ...      ...       ...       ...
     H2O  ...   0.0      0.0      0.0       NaN  0.000000
     H2O  ...   0.0      0.0      0.0       NaN  0.000000
     H2O  ...   0.0      0.0      0.0       NaN  0.000000
   C8H10  ...   7.0     36.0     38.0  3.222222  1.833333
   C8H10  ...   8.0     36.0     39.0  3.222222  1.861111
```

Figure 3. Data collected for the all the solvents with predefined HSPs.

Summary of Preliminary Data: After using Mordred to collect data about the parameters, the information was combined into one data file. This file allows for further manipulation with variable selection and training further in the project. Figure 1.2 shows part of the data file, with 1284 for the number of chemicals and 1889 columns representing the parameters received from Mordred.

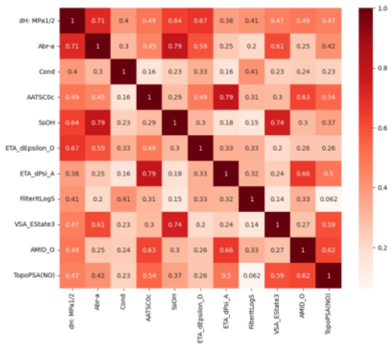
Expected Outcomes: The overall outcome of this aim is to collect as much and as high quality data as possible to have a robust training dataset. This data will be used in the machine learning model later in the project.

Potential Pitfalls and Alternative Strategies: If certain information is not available for a chemical, the try-except function allows the program to continue. Additionally, any rows with missing data are dropped before later steps, allowing the dataset to remain complete. However, missing data when calling the workflow may lead to an inability to calculate the best solvent mixture.

Specific Aim #2:

Following this step, these properties will be analyzed to determine which properties most affect the target Hansen Solubility parameter. This process includes an initial pass using Pearson and Spearman correlation coefficients, followed by further analyzation with the Boruta API. Boruta compares the parameters against the target parameter to decide which parameters most affect the target one. Finally, Pearson coefficients are used to determine any strong correlations between the two parameters. In this case, only one parameter is necessary to use in training as a high correlation indicates that the two parameters are conveying the same information. The smaller selected pool of parameters can be used in model training.

Justification and Feasibility: This step is important to understand what parameters are necessary for training. By limiting the number of parameters required, the model becomes more accessible to substances that may have less information available.



	dH: MPa1/2	Abr-a	Cond	AATSC0c	SsOH	ETA_dEpsilon_D	ETA_dPsi_A	FilterItLogS	VSA_EState3	AMID_O	TopoPSA(NO)
dH: MPa1/2	1	0.71	0.4	0.49	0.64	0.67	0.38	0.41	0.47	0.49	0.47
Abr-a	0.71	1	0.3	0.45	0.79	0.59	0.25	0.2	0.61	0.25	0.42
Cond	0.4	0.3	1	0.16	0.23	0.33	0.16	0.41	0.23	0.24	0.23
AATSC0c	0.49	0.45	0.16	1	0.29	0.49	0.79	0.31	0.3	0.63	0.54
SsOH	0.64	0.79	0.23	0.29	1	0.3	0.18	0.15	0.74	0.3	0.37
ETA_dEpsilon_D	0.67	0.59	0.33	0.49	0.3	1	0.33	0.33	0.2	0.26	0.26
ETA_dPsi_A	0.38	0.25	0.16	0.79	0.18	0.33	1	0.32	0.24	0.66	0.5
FilterItLogS	0.41	0.2	0.41	0.31	0.15	0.33	0.32	1	0.14	0.33	0.062
VSA_EState3	0.47	0.61	0.23	0.3	0.74	0.2	0.24	0.14	1	0.27	0.59
AMID_O	0.49	0.25	0.24	0.63	0.3	0.26	0.66	0.33	0.27	1	0.62
TopoPSA(NO)	0.47	0.42	0.23	0.54	0.37	0.26	0.5	0.062	0.59	0.62	1

Figure 4. Correlations of selected features against each other for hydrogen bonding HSP.

Summary of Preliminary Data: The correlation plot shown is an example plot for the dispersion parameter after selection using correlation coefficients and Boruta. Then, any variables with a correlation of 0.8 or greater were down-selected to one variable, as they conveyed similar information. The final parameters are represented here.

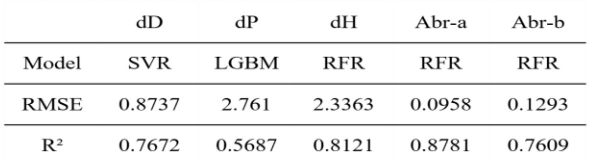
Expected Outcomes: This step should provide ample detailing of what parameters are most important for the prediction of HSPs.

Potential Pitfalls and Alternative Strategies: Given that Boruta can only take in a certain amount of data, the pre-selection process using multiple correlation factors allows a smaller subset to enter the process. Additionally, if there are too few parameters present at the end, certain steps in the process can be removed or altered to preserve data.

Specific Aim #3:

To understand which green solvents will dissolve the target polymers, solubility parameters must be compared. However, solubility parameters are not defined for a majority of green solvents; thus, this step looks at predicting HSPs through machine learning. First, a baseline program known as Lazy Predict will be used. This API allows a user to see the results of many different models quickly to determine which models are best for the data. Then, based on this data, a model will be fine-tuned to predict the parameters.

Justification and Feasibility: This step is necessary to understand which green solvents will work. Additionally, using a preliminary model test allows for a more efficient process as many models can be tested at once.



	dD	dP	dH	Abr-a	Abr-b
Model	SVR	LGBM	RFR	RFR	RFR
RMSE	0.8737	2.761	2.3363	0.0958	0.1293
R^2	0.7672	0.5687	0.8121	0.8781	0.7609

Figure 5. Initial model results for HSPs and Abraham parameters.

Summary of Preliminary Data: The R^2 coefficients and RMSEs were assessed for each of the three HSPs using their top predicted models from LazyPredict. Two other parameters, known as Abraham acidity and basicity parameters, were also included in this step, in case they proved useful in the prediction of solvents later on.

Expected Outcomes: The expected outcome of this step is the HSPs for every green solvent. Then, the HSPs of solvents and solvent blends can be compared to the target polymer's HSPs to determine the best green solvent.

Potential Pitfalls and Alternative Strategies: If these predictions are inaccurate, then alternative parameters can be included, such as Abraham parameters. Additionally, step 2 can be revisited to better decide which features are the most important to refine model outputs.

Specific Aim #4:

To validate the results of the model, it is necessary to test the solvent outputs in real life. This aim includes using the actual predicted solvents and using specific techniques to determine whether or not the solvent selection was successful. This step will require many trials, which leads to a need for a more automated approach for testing. Thus, a 3D printer and pump are modified to be used as a solvent dispenser to efficiently distribute specific solvent blends across a well plate.

Justification and Feasibility: This step is necessary to understand whether the model is successful. Using the modified device is also important to help eliminate human error and make it quicker to replicate the trials many times.

Figure 6. Percent error of pump before and after adjusting for constant correction factor.

Summary of Preliminary Data: This graph shows how the percent error in the pump was reduced greatly after offsetting the pump by a constant amount.

Expected Outcomes: The expected outcome of this stage is that the predicted solvents dissolve the target polymers.

Potential Pitfalls and Alternative Strategies: If the pump machine does not end up producing accurate data, dosing can be handled manually. Additionally, if the predicted solvents do not end up dissolving the target polymers, then more refinement of models and research on other important factors to consider during dissolution will occur. This strategy could include incorporating other parameters into the model such as Abraham acidity and basicity parameters.

Section III: Resources/Equipment

The main equipment for this project includes a variety of Python APIs. For physical equipment, the 3D printer-pump mechanism is modified during the project as a way to improve efficiency in dosing solvents. The solvents themselves are also important resources; however, given that only green solvents are considered, they are not toxic or harmful to work with.

Section V: Ethical Considerations

Ethically, this project helps to improve the condition of our planet as it reduces the need for production of plastic as well as replaces toxic solvents in the recycling industry. No organisms are being harmed during this experiment, nor are any people exposed to dangerous materials.

Section VI: Timeline

The timeline for this project is that specific aim one will be completed in October, specific aim two in December, specific aim three in January, and specific aim four in February. These times will likely change as the project progresses.

Section VII: Appendix

Section VIII: References

Abbot, S. (2025). *Hansen solubility parameters.* https://www.hansen-solubility.com/

Adam, H., Yousfi, M., Maazouz, A., & Lamnawar, K. (2025). Recycling of multilayer polymeric barrier films: An overview of recent pioneering works and main challenges. *Macromolecular Materials and Engineering*, *310*(7). https://doi.org/10.1002/mame.202400414

Al-Sakkari, E. G., Ragab, A., Amer, M., Ajao, O., Benali, M., Boffito, D. C., Dagdougui, H., & Amazouz, M. (2025). Ensemble machine learning to accelerate industrial decarbonization: Prediction of Hansen solubility parameters for streamlined chemical solvent selection. *Digital Chemical Engineering*, 14, 100207. https://doi.org/10.1016/j.dche.2024.100207

Ikegwu, U., del Carmen Munguía-López, A., Van Lehn, R. C., & Zavala, V. M. (2025). *Screening Green Solvents for Multilayer Plastic Film Recycling Processes.* ChemRxiv. https://doi.org/10.26434/chemrxiv-2025-zrmnp

Jiang, B., Yu, J., & Liu, Y. (2020). The environmental impact of plastic waste. *Journal of Environmental Earth Sciences*, *2*(2), 26–35. https://doi.org/10.30564/jees.v2i2.2340

Moriwaki, H., Tian, Y.-S., Kawashita, N., & Takagi, T. (2018). Mordred: A molecular descriptor calculator. *Journal of Cheminformatics, 10*(1). https://doi.org/10.1186/s13321-018-0258-y

Shershneva, E. G. (2021). Plastic waste: Global impact and ways to reduce environmental harm. *IOP Conference Series: Materials Science and Engineering*, *1079*(6), 062047. https://doi.org/10.1088/1757-899x/1079/6/062047

Tartakowski, Z. (2010). Recycling of packaging multilayer films: New Materials for Technical Products. *Resources, Conservation and Recycling*, *55*(2), 167–170. https://doi.org/10.1016/j.resconrec.2010.09.004

Toluene - overview. Occupational Safety and Health Administration. (2015, January 31). https://www.osha.gov/toluene

Vollmer, I., Jenks, M. J. F., Roelands, M. C. P., White, R. J., van Harmelen, T., de Wild, P., van der Laan,

G. P., Meirer, F., Keurentjes, J. T. F., & Weckhuysen, B. M. (2020). Beyond mechanical recycling: Giving new life to plastic waste. *Angewandte Chemie – International Edition, 59*(36), 15402–15423. https://doi.org/10.1002/anie.201915651

Walker, T. W., Frelka, N., Shen, Z., Chew, A. K., Banick, J., Grey, S., Kim, M. S., Dumesic, J. A., Van Lehn, R. C., & Huber, G. W. (2020). Recycling of multilayer plastic packaging materials by solvent-targeted recovery and precipitation. *Science advances, 6*(47). https://doi.org/10.1126/sciadv.aba7599

Winterton, N. (2021). The Green Solvent: A Critical Perspective. *Clean Technologies and Environmental Policy*, *23*(9), 2499–2522. https://doi.org/10.1007/s10098-021-02188-8

Zhou, A., Longacre, L., Motati, R., & Acree, W. E. (2023). Abraham Solvation parameter model: Revised predictive expressions for solute transfer into polydimethylsiloxane based on much larger and chemically diverse datasets. *Compounds, 3(*1), 205–223. https://doi.org/10.3390/compounds3010017