A machine learning approach for modeling and inverse design of polymer blends

Soham Jariwala[1], Claudio César Claros Olivares[2], Zijie Wu[1]
[1] Department of Chemical and Biomolecular Engineering
[2] Department of Electrical and Computer Engineering

Mentor: Dr. Nicolae C. Iovanac
DuPont de Nemours, Inc.

Spring, 2022

CHEG/CISC 867
Computing and Data Science for Soft Materials Innovation & Discovery
Instructors: Prof. Arthi Jayaraman, Prof. Sunita Chandrasekaran

Abstract

1 Introduction

Polymers have hierarchical structures and complex interactions that span several orders of magnitude in length and time scales. Changing the composition and processing conditions gives rise to tunable physico-chemical properties that can be engineered to obtain a desired physical, chemical and electrical functionality. For a polymer with well defined composition and processing conditions, one may be able to apply simulation techniques, such as Monte Carlo (MC), molecular dynamics (MD), and density functional theory (DFT) to obtain useful quantitative relationships for structure, composition, and physico-chemical properties[1].

However, in the case of polymer blends, where several ingredients (small molecules or polymers) are added to the base polymer to improve its functionality and performance, obtaining a quantitative relationship from simulations may become computationally intractable. The computational cost of adequately capturing the interactions as well as exploring the complete phase space with several constituents becomes quite enormous, especially considering the number of computations required to enable the inverse design. As a results, formulation design of polymer blends relies heavily on heuristics and empirical relationships. This alternative approach is also called polymer informatics[2, 3], where machine-learning methods and data-driven modeling can accelerate the discovery and development of new formulation, allowing one to screen candidate polymer formulations for desired target properties.

In this work, we apply a polymer informatics approach to predict the following three target properties from a polymer formulation dataset generated by DuPont,

- *Water absorption*: It is a quantitative measure of water absorbed by the polymer blend, usually reported as a percent water per dry weight. In addition to the type of polymers and additives used, temperature and exposure time also have to be considered[4].

- *Hardness*: It is defined as resistance to local plastic deformation. The tests are usually performed by measuring the depth of indentation against an applied load. Values are reported as dimensionless number on scales that are specific to the type of test conducted. Barcol, Rockwell (E,M,R), and Shore tests are some of the hardness tests suitable for polymers. Hardness is strongly influenced by the presence of additives and testing temperature[5].

- *Thermal conductivity*: It is the ability of the polymer blend to conduct heat reported in terms of heat flux per temperature gradient (units W m $^{-1}$ K $^{-1}$). Thermal conductivity depends on the nanoscale structure of the polymer chains and additives as well as the presence of crystalline and amorphous domains. Processing plays a key role in determining the thermal conductivity of the end product[6].

2 Project Goal

The goal of this project is to connect the aforementioned target properties to the ingredients of the polymer formulation. The polymer formulation dataset comprises of 79 distinct formulations containing ingredients such as plasticizers, toughenners, antioxidants, flame retardants and stabilizers in addition to the base polymer and fillers. The dataset is aliased to protect the intellectual property of DuPont, therefore, the identities and structural characteristics of these ingredients are not included in the model. Our approach can be broken down into the following steps:

- *Forward design*: A machine-learning (ML) based regression model is systematically selected and trained for each of the three target properties using 71 of the 79 formulations. The hyper-parameters are tuned using cross-validation and model predictions compared against the eight held-out formulations.
- *Inverse design*: A particle swarm optimization scheme is implemented to predict the polymer formulations from the ML-based regression models generated in the previous stage.

References

[1] T. E. Gartner and A. Jayaraman, "Modeling and simulations of polymers: A roadmap," *Macromolecules*, vol. 52, no. 3, pp. 755–786, Feb. 2019. DOI: 10.1021/acs.macromol.8b01836.

[2] D. J. Audus and J. J. de Pablo, "Polymer informatics: Opportunities and challenges," *ACS Macro Letters*, vol. 6, no. 10, pp. 1078–1082, Oct. 2017. DOI: 10.1021/acsmacrolett.7b00228.

[3] L. Chen *et al.*, "Polymer informatics: Current status and critical next steps," *Materials Science and Engineering: R: Reports*, vol. 144, p. 100 595, 2021. DOI: https://doi.org/10.1016/j.mser.2020.100595.

[4] G. Baschek, G. Hartwig, and F. Zahradnik, "Effect of water absorption in polymers at low and high temperatures," *Polymer*, vol. 40, no. 12, pp. 3433–3441, 1999. DOI: https://doi.org/10.1016/S0032-3861(98)00560-6.

[5] polymerdatabase.com. "Hardness of plastics and rubbers." (2022), [Online]. Available: https://polymerdatabase.com/polymer%20physics/ShoreRockwell.html.

[6] C. Huang, X. Qian, and R. Yang, "Thermal conductivity of polymers and polymer nanocomposites," *Materials Science and Engineering: R: Reports*, vol. 132, pp. 1–22, 2018. DOI: https://doi.org/10.1016/j.mser.2018.06.002.