A novel domain knowledge-informed machine learning approach for modeling solid waste management systems

Rui He and Mitchell J. Small

Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, United States

Ian Scott*

NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Campus de Campolide, 1070-312, Lisbon, Portugal

Motolani Olarinre

Department of Machine Learning, Carnegie Mellon University, Pittsburgh, PA 15213, United States

Mexitli Sandoval-Reyes and Paulo Ferrão

IN+ Center for Innovation, Technology and Policy Research, LARSyS, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal

(*) Correspondence author: iscott@novaims.unl.pt

He, R., Small, M. J., Scott, I. J., Olanrinre, M., Sandoval-Reyes, M., & Ferrão, P. (2023). A Novel Domain Knowledge-Informed Machine Learning Approach for Modeling Solid Waste Management Systems. Environmental Science & Technology, A-J. https://doi.org/10.1021/acs.est.3c04214

ABSTRACT

Sustainability challenges, such as solid waste management, are usually scientifically complex and data scarce, which makes them not amenable to science-based analytical forms nor data-intensive learning paradigms. Deep integration between data science and sustainability science in highly complementary manners offers new opportunities to tackle these conundrums. This study develops a novel hybrid neural network (HNN) model that imposes the holistic decision-making context of solid waste management systems (SWMS) on a traditional neural network (NN) architecture. Equipped with adaptable hybridization designs of hand-crafted model structure, constrained or predetermined parameters, and customized loss function, the HNN model is capable of learning various technical, economic, and social aspects of SWMS from a small and heterogeneous dataset. In comparison, the versatile HNN model not only outperforms traditional NN models in convergence rates, which leads to a 22% lower mean testing error of 0.20, but also offers superior interpretability. The HNN model is capable of generating insights into the enabling factors, policy interventions, and driving forces of SWMS, laying a solid foundation for data-driven decision-making.

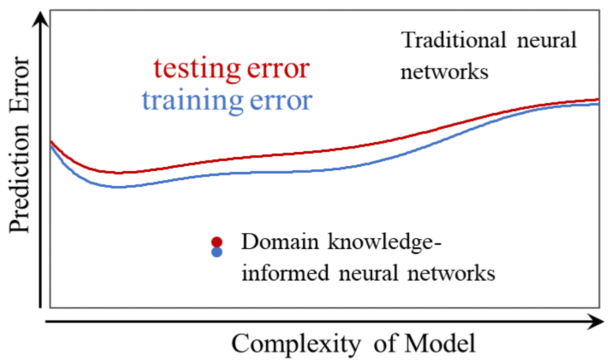
HIGHLIGHTS

- novel hybridization of sustainability science with data science
- system thinking of the decision-making context of SWMS
- versatile model for decision support and scientific discovery
- improved neural network performance on small datasets

- increased interpretability of neural network decisions

KEYWORDS: waste management, decision support, machine learning, hybrid neural network, physics-informed ML, interpretable ML

SYNOPSIS: A novel hybrid machine learning model was developed by “handcrafting” neural networks based on the domain knowledge of solid waste management for improvements in performance, interpretability, and decision support.



1. INTRODUCTION

Fueled by growing global waste generation,[1] worsening impacts of climate change, and tightening resource constraints, societies are beginning to pivot to a circular economy model, featuring loop-closing opportunities of redesign, reduce, reuse, and recovery along the life cycles of products and resources. Under this circular economy paradigm, solid waste management systems (SWMS), responsible for handling municipal solid waste (MSW) streams, need to serve the dual purposes of waste disposal and waste valorization in technologically feasible, economically beneficial, environmentally responsible, and socially benign ways. To support the design and operations of SWMS in this new era, decision support tools are instrumental to policy makers and waste managers steering these multi-stakeholder, nonlinear, complex systems onto a circular economy trajectory.[2–5]

Existing SWMS tools can be categorized into system engineering models that support the strategic and operational decision-making of SWMS (e.g., optimization models, multi-criteria decision-making, etc.) and system assessment tools that evaluate the performance of SWMS (such as life cycle assessment and material flow analysis).[3] The state-of-the-art system engineering models, such as Solid Waste Infrastructure Modelling System (SWIMS), can optimize regional SWMS configurations including the deployment of treatment technologies and allocation of MSW to these technologies based on both environmental and economic objectives and constraints.[6] The SWIMS model can also dynamically schedule future waste treatment capacity building based on different policy and waste generation scenarios. As a latest system assessment tool, Solid Waste Optimization Life-cycle Framework in Python (SwolfPy) offers an open-source, Python-based software with built-in Monte Carlo simulation and optimization features to dynamically assess the environmental and economic impacts of SWMS handling heterogeneous waste inputs.[7]

Powerful as SWMS decision support tools have become, challenges still persist when it comes to application. The main challenge resides in the disconnections between the calculated theoretical optimums and the status quo of existing SWMS. Many of these recommendations might never be adopted or implemented because they fail to overcome the inertia of existing systems. Moreover, system expansion is needed to integrate social, economic, institutional, and environmental aspects to comprehend the holistic decision-making context of SWMS.[3,4] Traditional decision support tools with a "tunnel vision" of the technical aspects, focusing on waste generation, logistics, and technologies, may lead to impractical or erroneous recommendations. For example, these tools might recommend waste-to-energy (WtE) technologies without factoring in the negative perception-induced public resistance,[8] or simply underestimate the future capacity demand by losing sight of the international trade restrictions such as the China's waste import bans.[9] Lastly, the model complexity, parameter variability, and data intensity jeopardize the generalizability and practicality of these models as decision support tools for policy makers.

Meanwhile, data-driven approaches such as machine learning (ML) models have gained tremendous momentum in recent years, due to their superior capability of simulating complex nonlinear system behaviors and the recent improvements in computational power.[10] However, the application of ML in SWMS decision support is still in the natal stage, restricted by the bottleneck issues of data availability and quality.[11–13] Recent studies resorting to graphical representations such as Bayesian belief networks,[14] fuzzy cognitive maps,[15] and system dynamics[16–19] aim to navigate around this issue by combining domain knowledge with empirical evidence. These models essentially either rely on existing scientific understandings to impose structures on data-driven models or harness insights from empirical data to bridge knowledge gaps, forging a new partnership between data science and sustainability science.[20,21]

Inspired by this hybrid problem solving paradigm, deep integration is proposed between science-based domain knowledge and data-driven ML models by leveraging the advantages of both in a highly complementary manner. Science-based rules and domain knowledge can complement ML models by confining their search spaces (i.e., reducing degrees of freedom), thus lowering data requirements and computational intensity. On the other hand, ML models can shed light on mechanisms that are too complex to unveil using traditional science-based methods but are embedded in the hidden patterns of data. This synergistic relationship can be analogized to “digging a tunnel from both sides” as depicted in Figure 1, where sustainability scientists and data scientists work collaboratively through an iterative process of learning from each other to adjust their efforts and directions.[22]

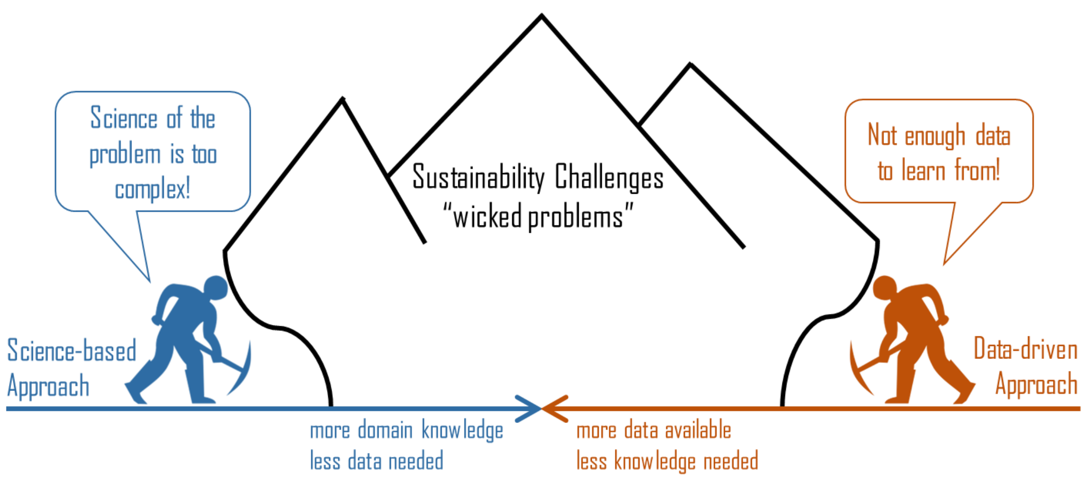


Figure 1 Concept of the hybrid modeling approach to tackle sustainability challenges

This study demonstrates that the domain knowledge of SWMS can be harnessed to constrain ML models to improve prediction performance and model interpretability. Specifically, a novel hybrid neural network model is developed as a first attempt to simulate various market, social, and technical mechanisms of SWMS and to identify policy levers that can mobilize social and financial capitals to expedite transitions to a circular economy. This model is intended to complement existing optimization tools, as it aims to offer roadmaps and action plans to reach the "ideal destinations" suggested by optimization models. The following sections of this paper describe the decision-making context of SWMS, the hybrid ML model development and results, and its potential applications in policy decision support.

2. METHODOLOGY

2.1 System perspective of the SWMS decision-making context

System thinking is crucial to unravel the complexity of the interconnected social, economic, governmental, and environmental dimensions of sustainability challenges.[23] This approach requires an expanded system boundary that encompasses various behavioral, financial, and institutional factors, in addition to technical considerations, as endogenous variables to reflect the complex decision-making context of SWMS.[24] To identify the key components of SWMS and their interconnections, extensive literature review is conducted on waste management systems and drivers,[15,25–27] the social factors,[28–30] the environmental impacts,[31–33] the market mechanisms,[4,27,34] and the technical treatment pathways of SWMS.[6,35,36]

Comprehensive lists of decision variables (policy and technical decisions that shape SWMS), state variables (variables that describe the state of SWMS), and driver variables (state variables that drive SWMS changes) are extracted during the literature review to build a comprehensive system perspective of SWMS. This systemic representation is subsequently simplified and transfigured into a graphical model by removing non-essential or unmeasured variables to improve computational feasibility and model interpretability. The definitions of all the variables, extracted and assumed relationships, and simplification rationales are detailed in Table S1 and S2 of the Supporting Information.

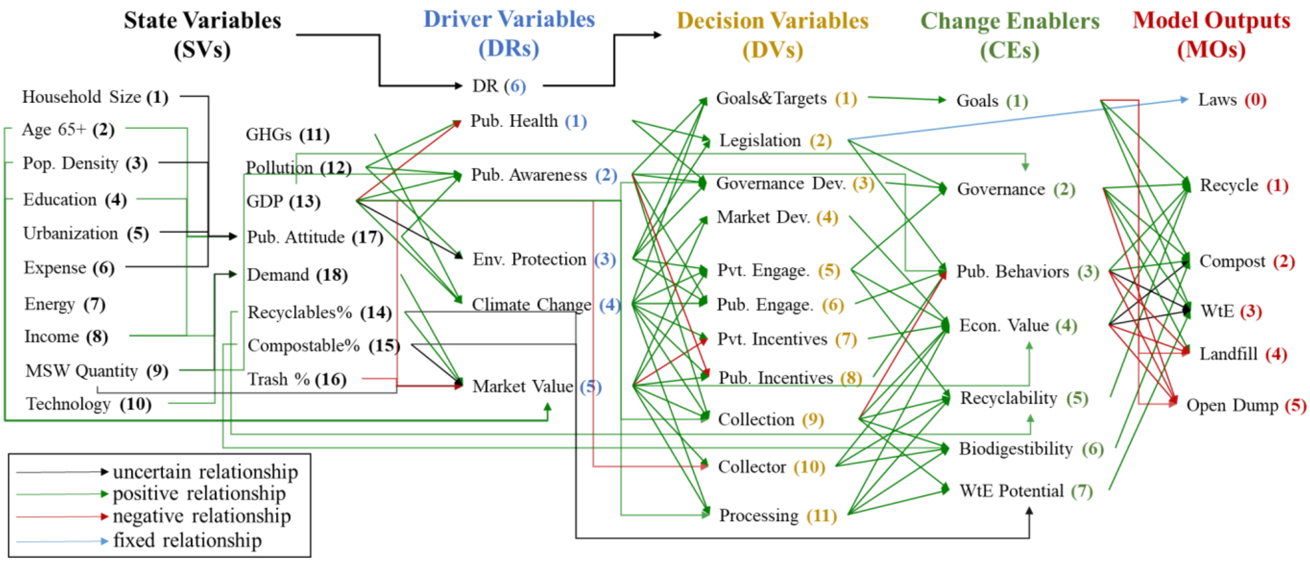


Figure 2 Simplified decision-making context of SWMS based on system thinking. Variables (nodes) are presented in 5 layers of SVs, DRs, DVs, CEs, and MOs. Relationships (directed edges) are constrained based on domain knowledge and assumptions to be deterministic (blue), positive (green), or negative (red). Relationships lacking domain knowledge are unconstrained (black).

As illustrated in Figure 2, the 18 state variables (SVs), describing socioeconomic, demographic, environmental, and waste generation conditions of SWMS, determine the power of 5 identified driver variables (DRs) of SWMS: *Public Health*, *Public Awareness*, *Environmental Protection*, *Climate Change*, and *Market Value*. To factor in the possibility of incomplete knowledge, an "unknown" driver of DR6 is added, which is fully connected to all the SVs and policy decisions. These driving forces can actuate 11 institutional, social, economic, and technical decisions (DVs) at the disposal of SWMS decision makers and managers. A layer of "change enablers" (CEs) is inserted between the DVs and the model outputs (MOs) to incorporate the feedback from policy interventions on these change enabling factors, such as sustainable public behaviors (CE3), higher economic value of MSW (CE4), increased technical feasibility (CE5-7), and expanded governance capacity (CE1-2). Finally, it is these CEs that facilitate different waste management practices of recycling, composting, waste-to-energy (WtE), landfilling, and open dumping. Other waste management options, such as reuse, anaerobic digestion, advanced WtE technologies are out of the scope of this study due to poor data availability.

Relationships among the variables, displayed as directed edges in Figure 2, are either qualitatively extracted from literature review or hypothesized based on discretion of the authors (detailed in Table S1 of the Supporting Information). For example, it is believed that *Public Health* (DR1) remains an important driver for developing countries and leads to legislation and higher management capacity.[24,27,37] Thus, it receives a negative input from *GDP* (SV13) and feeds positively to *Legislation* (DV2) and *Governance Development* (DV3), which represents governmental waste management capacity building. In addition, the *Public Health* driver is also assumed to be positively correlated with higher levels of *Pollution* (SV12).

The system perspective in Figure 2 integrates both technical mechanisms and non-technical aspects of MSW management into a holistic view of the SWMS decision-making context. Technical mechanisms are embedded in the technical enabling factors (CE5-7) determined by waste collection systems, waste processing technologies, and MSW compositions. Various nodes and edges connected to the *Economic Value* (CE4), *Public Behaviors* (CE3), *Goals* (CE1), and *Governance* (CE2) represent the market, behavioral, and institutional forces that shape SWMS. It is worth clarifying that Figure 2 is not intended to be a complete and robust representation of SWMS. Instead, it is more indicative of deliberately pooled initial understandings of SWMS researchers as an inception of the hybrid modeling approach, which is designed to refine SWMS knowledge through an iterative process.

2.2 Hybrid machine learning model based on domain knowledge

The complex, nonlinear relationships in Figure 2, most of which are not amenable to explicit analytical forms, can be empirically simulated with multilayer feedforward neural networks (NNs).[38] As a universal function approximator,[39] NN is a powerful computation tool that has been successfully applied to predict MSW generation,[40] properties,[41] and management.[13] In a typical feedforward NN model with multiple hidden layers, input data propagate through the entire network after certain linear combination and nonlinear (sigmoid) activation at each node (or neuron) in the subsequent layers. A novel hybrid neural network (HNN) model is proposed by hand-building a NN based on the same structure as the directed acyclic graph of Figure 2. The HNN model can be mathematically described as:

$$Node_j = tanh\left(x_j\right) = \frac{\exp(x_j) - \exp(-x_j)}{\exp(x_j) + \exp(-x_j)}\ (for\ Node_j\ not\ in\ the\ output\ layer) \quad \text{eq. 1a}$$

$$Node_j = logistic\left(x_j\right) = \frac{1}{(1 + \exp(-x_j))}\ (for\ Node_j\ in\ the\ output\ layer) \quad \text{eq. 1b}$$

$$x_j = \sum_i^{all\ the\ inputs\ of\ Node_j} (Node_i \times \alpha_{ij}) + bias_j \quad \text{eq. 2}$$

$$\boldsymbol{\alpha}', \boldsymbol{bias}' = \underset{\boldsymbol{\alpha},\boldsymbol{bias}}{\operatorname{argmin}} J = \underset{\boldsymbol{\alpha},\boldsymbol{bias}}{\operatorname{argmin}} \frac{1}{N} \sum_{p=1}^{N} \sum_{j=0}^{5} (MO_{pj} - \widehat{MO_{pj}})^2 \quad \text{eq. 3}$$

where: $Node_j$ is a node with input node(s). $\alpha_{ij}\ and\ bias_j$ are weight and bias parameters (unitless coefficients). MO_{pj} is the observed value of the j^{th} output node for the p^{th} observation. $\widehat{MO_{pj}}$ is the model simulated value of the j^{th} output node for the p^{th} observation. N is the number of observations in the dataset, and J is the quadratic loss function for evaluating prediction errors (mean squared errors). $\boldsymbol{\alpha}', \boldsymbol{bias}'$ are the learnt parameters which minimize the loss function J.

By imposing the decision-making context of SWMS on an NN architecture, degrees of freedom of the search space can be dramatically reduced as compared to a typical fully connected NN, enabling this "handcrafted" HNN to achieve relatively satisfactory prediction performance with a limited number of training cases. In addition to the "hand-crafted" neurons and edges, five other hybridization treatments are devised and implemented to further constrain the HNN model with domain knowledge, making it a so-called "informed machine learning model".[42] These treatments are variations and integration of different physics-informed model designs reported in recent ML literatures,[43–45] with which the HNN approach should be applicable to any systems that can be expressed as differentiable computation graphs (acyclic directed graphs expressing equational data).[38] These five hybridization treatments are listed below.

1. Weight of parameters (α_{ij}) can be constrained to be positive, negative, or fixed based on domain knowledge (e.g., all the red, green, and blue edges in Figure 2).
2. Neurons (the forms of eq. 1a, 1b, and 2) can be predetermined based on domain knowledge.
3. Neurons can feed to multiple layers, as long as the network is acyclic (e.g., SV13, SV14, DR2, etc.)
4. Neurons with empirical observations in the middle layers can be added to the loss function to be trained simultaneously with the model outputs (e.g., DV2 is constrained by training on MO0, which equates DV2).
5. Feedback loops can be incorporated by inserting and connecting neurons on different layers that share the same physical connotation (e.g., DR5 and CE4).

In terms of model outputs, the HNN model not only predicts the probability of each final treatment option, but also indicates the influential policy decisions and socioeconomic drivers that lead to these predictions. However, this superior interpretability is predicated upon the superimpositions of physical connotations on neurons, which should be treated as “pseudo-physical” before they can be verified by science or supported by data. Because it is presumptuous to use simple sigmoid transformation of linear combinations of variables to model the entire SWMS. Nonetheless, the HNN model in its current configuration does provide a generic starting point for iterative future improvements once more robust domain knowledge or data is acquired. More discussions on the validity and implementation of these hybridization treatments are provided in the Supporting Information.

2.3 Data collection and preparation

Most input variables (SV1-SV7, SV9, SV13-SV16) and all the output variables (MO0-MO5) of the HNN model are extracted from an existing SWMS knowledge base.[46] Additional inputs of GNI per capita (SV8), CO_2 emissions (SV11), and PM2.5 pollution (SV12) are collected from the World Bank Database.[47] Technology innovation indexes (SV10) are taken from the “Global Innovation Index 2021” report.[48] Input variables of *MSW Quantity*, *Population Density*, *Household Size*, *Energy*, *GDP*, *Income*, *GHGs*, and *Pollution* are logarithmically transformed for conformation to near normality. A final dataset with a total of 1,376 observations (cases) is compiled after removal of missing data (except for SV14-16, where missing data are replaced with average values) and normalization (to rescale input variables into a range of 0 to 1). Finally, 278 observations (or roughly 20% of the entire dataset) are randomly selected and held out for model testing purposes (the testing set). The rest of the data (or 80%) are used for 5-fold cross-validation (CV), with training sets of 878 cases and validation sets of 220 cases for model evaluation, hyperparameters selection, and subsequently, model training. Detailed information on input data and data preparation is provided in Section 1.2 of the Supporting Information.

With a mixture of country-level and city-level observations, biased sampling towards developed countries (due to data availability), multiple data sources, and prevalence of missing data, this small, heterogeneous dataset poses a challenging training scenario that sustainability researchers typically face.

2.4 Model parameterization, training, and evaluation

Parameters of the HNN model (α_{ij} and $bias_j$) are determined by training with backpropagation using various gradient descent algorithms (built in the Python package of PyTorch).[49] All the parameters are randomly initialized from a normal distribution with a mean of 0 and a standard deviation of 0.5. To alleviate the problems of overfitting and overtrained parameters, an L2 regularization term is added to the loss function J.

$$J' = \frac{1}{N}\sum_{p=1}^{N}\sum_{j=0}^{5}(MO_{pj} - \widehat{MO_{pj}})^2 + \gamma\sum_{\beta=all\ \alpha\ and\ bias}\beta^2 \quad \text{eq. 4}$$

where: γ is the weight of the L2 regularization term, β represents all the model parameters.

Before the final model training and evaluation, an extensive grid search is performed with a 5-fold CV on the remaining 80% of data, tuning various hyperparameters of the model, including: activation functions, learning rates, regularization methods, gradient descent algorithms, and training iterations (epochs). In addition, to leverage the knowledge discovery capability of ML, the model structure depicted in Figure 2 is relaxed to explore similar HNNs with and without DR6.

Finally, the model is trained with the hyperparameters that minimize the prediction errors (MSEs) on validation sets of the 5-fold CV and evaluated of J on the unobserved testing set. If multiple sets of weights minimize J on the testing set, the interpretability of these learnt weights is then used as the main criteria for determining the final HNN model outcomes.

To evaluate the performance of the HNN model, traditional fully connected NN models with various degrees of freedom are built by varying the number of neurons in 2 hidden layers. Both HNN and NN models are trained on the same training data (1,098 cases) and then evaluated on the same testing set (278 cases). The NN models share the same hyperparameters as the HNN model for fair comparisons.

Data processing is conducted in the Microsoft Excel software. The HNN model and the traditional NN models are built, trained, and tested in the Python 3.8.16 environment. The Python code with the input data files is provided in the Supporting Information.

3. RESULTS AND DISCUSSIONS

3.1 Results of model training and selection

Based on the hyperparameter tuning exercise, L2 regularization (γ =0.000001), adagrad optimization approach, a learning rate of 0.3, and a training time of 100 epochs were selected for the HNN model training. While the “tanh” activation function (eq. 1a) performed the best, it complicated the interpretability of the model outputs. Thus, it was replaced by the logistic activation function (eq. 1b) only for the output layer (MOs), which helps constrain model outputs within the range of 0 to 1. Details and results of this exercise can be found in the Supporting Information.

Adagrad outperformed other gradient descent algorithms in its consistency in reaching the lowest training and testing errors. Adagrad dynamically adjusted individual gradient for each parameter based on the previous steps, which reduced the chance of landing at local minima of non-convex search space. L2 regularization term penalized large (overtrained) weights and was instrumental in generating more evenly distributed weights, which also helped with model interpretability. In addition, a relatively short training time of 100 epochs was chosen in conjunction with a moderate learning rate of 0.3 to control overfitting and overtraining.

One unique design of the HNN is the constrained model weights, which complicates the non-convex search space and may lead to sub-optimal training results. To ease this concern, an explorative experiment was conducted by removing all the constraints on model parameters (i.e., treating all the red and green arrows as black arrows in Figure 2). The resulting model training errors and testing errors are not statistically different from those achieved by the HNN with constrained parameters. However, the resulting weights are mostly similar in magnitude but different in +/- signs. This experiment verifies that the superior interpretability of HNN model does not compromise the goodness-of-fit.

As Figure 2 was treated as preliminary understandings of SWMS, different variations of it were built and evaluated to finalize the model structure. These variations mainly involve the addition of DR6 (the “unknown” driver variable) and the connections between *GDP* and DV3, DV9, DV10, and DV11. The results revealed that adding DR6 significantly improved the model performance, while adding connections between *GDP* and DVs helped generate more interpretable model weights. Thus, the final model structure incorporated DR6 and these connections as presented in Figure 2.

3.2 The final HNN model and validation

After 100 epochs of training, the prediction error (*J*) on the training data consistently dropped from above 0.300 to around 0.215. All the 143 model parameters (not including biases) are presented in Figure 3 in the same fashion as Figure 2 with added thickness of arrows representing the value of the learnt weights. The complete numerical results are provided in Table S4 of the Supporting Information.

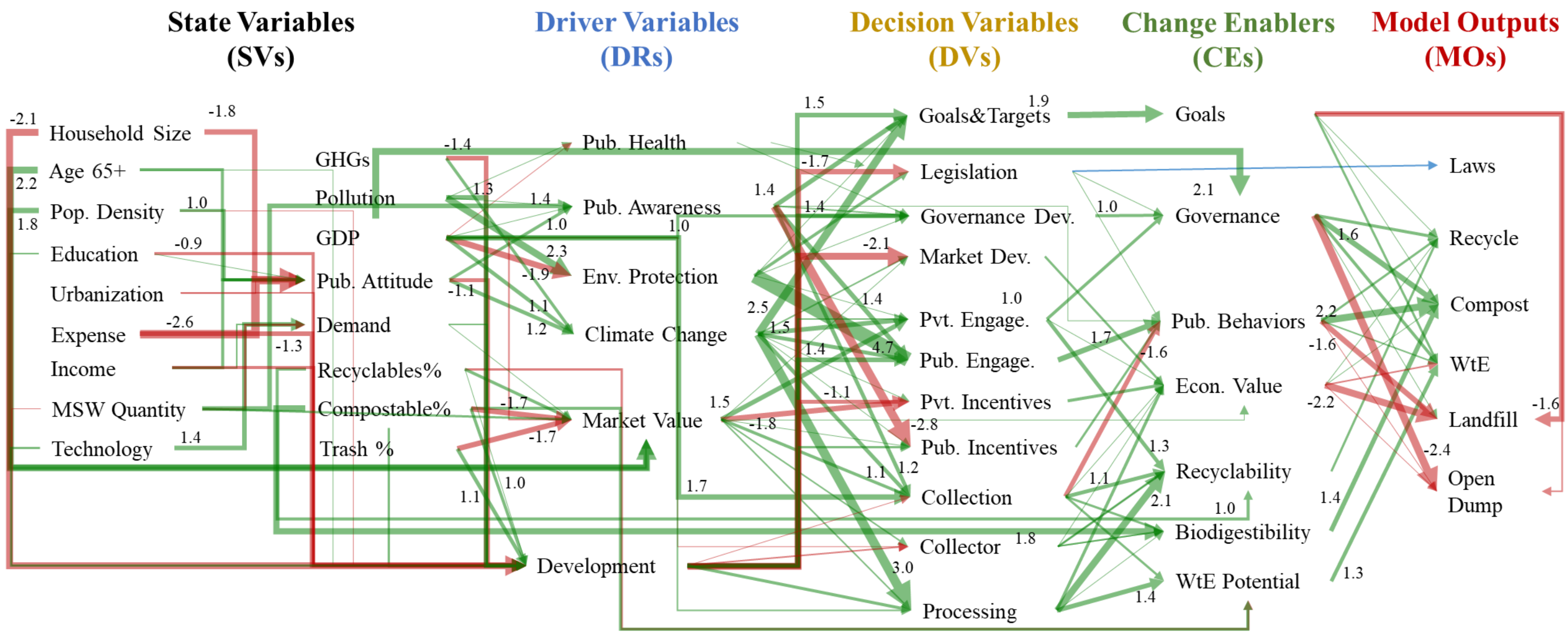


Figure 3 The learnt weights of the HNN model. Red arrows represent negative weights and green arrows represent positive weights, with the thickness of arrows and the numbers (above "1" or below "-1") next to the arrows representing the numerical value of these weights. DR6 in Figure 2 is interpreted and presented as "*Development*".

One way to verify the validity of the trained HNN model is to interpret the unconstrained model parameters. For example, *Public Attitude* (SV17) turns out to be negatively associated with *Household Size* (-1.8) and household *Expense* (-2.6), while positively associated with *Population Density* (1.0) in addition to the constrained parameters of *Age 65+* (0.90), *Education* (0.35), and *Income* (0.87). Most of the unconstrained parameters are associated with DR6. By analyzing the significant inputs of *Household Size* (-2.1), *Education* (-0.91), *Expense* (-1.3), *Income* (-0.89), Technology (0.8), *Pollution* (1.4), *GDP* (0.80), and *Public Attitude* (-1.1), it can be inferred that a high value of DR6 corresponds to a relatively low level of socioeconomic development (but good growth momentum). Thus, the DR6 in Figure 2 can be interpreted as a new driver of "*Development*". Furthermore, its connections with DVs indicate that for countries at relatively low socioeconomic development levels with good growth momentum, *Governance Development* (capacity building), *Public Engagement* (education, decision involvement, etc.), and *Processing* (upgrading waste processing technology) are among the most common policy interventions, while pursuing market development, reinforcing laws, and providing private incentives are among the least adopted.

Although most parameters are constrained, their absolute values can shed light on the relative importance of their corresponding input variables. For example, *Public Awareness* is largely shaped by *Public Attitude* (1.0) and *MSW Quantity* (1.4), as compared to *Pollution* (0.29) and *GDP* (0.34). Responsible *Public Behavior* can be attributed to more direct *Public Engagement* (1.7) and *Public Incentives* (0.83), or undermined by MSW *Collection* system upgrades (-1.6) since multi-stream collection tends to compromise the convenience of use by the public.[50] On the other hand, less direct interventions such as environmental *Legislation* (0.09) and higher *Public Awareness* (0.03) seem to be less influential. The *Economic Value* of MSW receives weighty boost from

Market Development (0.72), *Private Incentives* (0.93), and waste *Processing* technology investment (0.88). These weights signal the determining role of infrastructure investments and financial incentives in creating favorable market conditions for MSW management. In terms of technical feasibility, *Recyclability*, *Biodigestibility*, and *WtE Potential* are mainly empowered by advanced collection systems and processing technologies, while being highly sensitive to the composition of MSW. Overall, these weights suggest that the HNN model is able to make sense of certain technical, social, and economic mechanisms of SWMS with the help of domain knowledge constraints.

On the other hand, weights close to zero imply insignificant contributions, such as those for *Public Health* and the influence of *Market Value* on *Economic Value*. However, these weights do not necessarily nullify the assumed importance of these factors in SWMS decision-making, but rather indicate an unsuccessful modeling attempt using the training dataset. This outcome is due to three possible reasons. First, these factors may not have much influence on the model outputs indeed. Second, it is possible that the small, noisy, heterogeneous dataset does not allow learning of these factors. In other words, the influences of *Public Health* may be masked by the noise of data, or not captured by the dataset in the first place given that the majority of observations are drawn from developed countries. Third, it is possible that these factors are modeled with insufficient model complexity or insufficient data inputs.

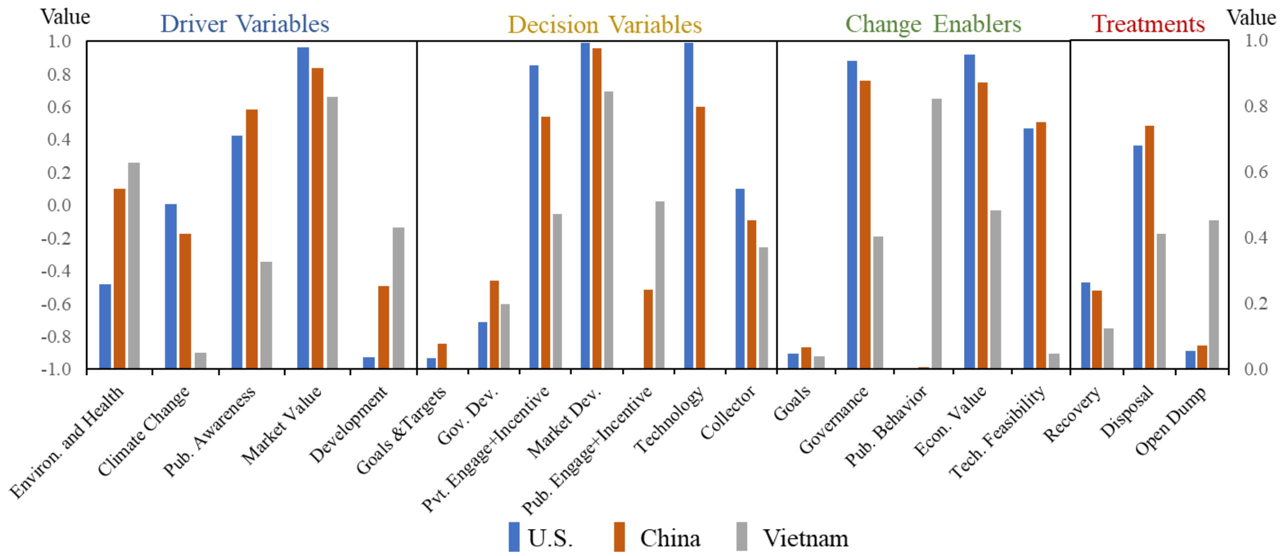


Figure 4 Simulated values of DRs, DVs, CEs, and MOs of the U.S., China, and Vietnam during historical period (circa 2015). All variables are constrained within the range of -1 to 1 except for "Treatments". "Technology" is the sum of DV9 and DV11, "Tech. Feasibility" is the average of CE5-7. "Recovery" is the sum of *Recycle* and *Compost*, and "Disposal" is the sum of *WtE* and *Landfill*.

To further demonstrate model interpretability and validity, model simulations of DRs, DVs, CEs, and MOs were made for the U.S., China, and Vietnam as testing cases, respectively. As shown in Figure 4, the HNN model not only captures the differences in waste treatment practices among these countries, but also uncovers the driving forces and enabling factors leading to these outcomes. It is worth noting that all the variables are unitless and constrained between -1 and 1 or 0 and 1 due to the form of activation functions (eq. 1a and 1b). Values are comparable across different countries, but not across variables.

It is suggested that *Market Value* of MSW, *Public Awareness*, and *Climate Change* are more prominent drivers of SWMS improvements in the U.S., while *Environmental Protection*, *Public*

Health, and socioeconomic *Development* agendas fuel the MSW management in Vietnam. China is propelled by all 6 drivers, especially high public awareness, given its complex socioeconomic and environmental conditions, as well as a strong booming economy. These different drivers shape U.S. SWMS that feature more aggressive private sector engagement, technology investments, and market development. In less developed economies, such as Vietnam, SWMS decision makers tend to resort to financially incentivizing the public sector, educating the general public, and engaging informal waste pickers in MSW collection. In the U.S., MSW treatment features value recovery practices enabled by technological proficiency, high economic value, and high governance capacity. On the other hand, with limited access to technologies and waste management capacity, SWMS in Vietnam are mainly reliant on public behaviors, which leads to a higher percentage of open dumping.

3.3 Evaluation of the hybrid modeling approach

Oftentimes, researchers of sustainability systems find themselves stranded by the barriers of overwhelmingly complex science and underwhelmingly scarce data. The hybrid problem-solving paradigm combining sustainability science with data science offers new opportunities to break through such conundrums. The HNN model combines the interpretability of a parametric model by parameterizing implicit relationships with sigmoid functions and the flexibility of a non-parametric model by being free from distribution assumptions. To demonstrate the advantages of the HNN model, traditional fully connected NN models were trained and tested in the same way as the HNN model. Their prediction errors (J) on the training data (1,098 cases) and the testing set (278 cases) are contrasted in Figure 5.

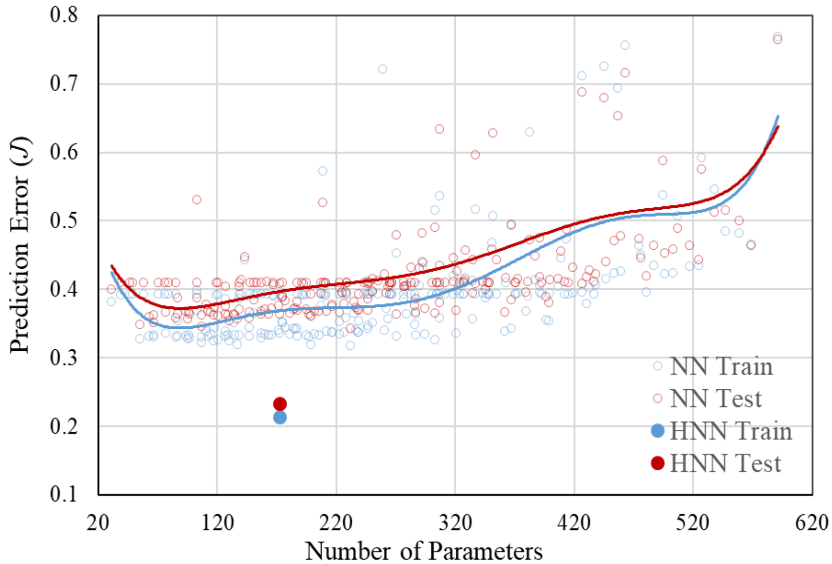


Figure 5 Performance comparisons between the HNN and NN models after 100 epochs of training. The two curves represent the mean squared errors (MSEs) of model outputs (J) evaluated on the training (blue) and testing (red) data of the traditional NN models, while the two dots represent these errors of the HNN model.

As illustrated in Figure 5, the "handcrafted" HNN model outperforms traditional NN models in the training stage with a significantly lower training error of 0.215, thanks to a much faster convergence rate. On the testing data, the HNN model yields a significantly lower error (0.234) than all the NN models receiving 100 iterations of training. The traditional NN models with more parameters yield higher prediction errors mainly because of the short training time. Given sufficient model complexity (over 1,500 parameters) and training time (30,000 epochs), the traditional NN models reach around a minimum testing error of 0.164, which can be deemed as the theoretical limits of learning for the datasets in this study. This suggests that the HNN model with merely 100 training iterations yields an average testing error ($\sqrt{J/6}$ or RMSE) of 0.197, which is 22% lower than that of the NN models, and only 0.032 higher than the theoretical limit.

Furthermore, erratic learning behaviors were observed during the training of NN models, implying challenges of NN models in learning from small datasets with high noise levels. In sum, the HNN model excels in convergence rate and outperforms the traditional NN models with similar amounts of training in prediction accuracy.

Adopting this science-ML hybrid approach offers flexibility to cope with three common research challenges. First, in cases of missing data, additional layers can be appended to approximate complete input data by applying feature learning techniques that are commonly adopted in ML.[51] Second, with large numbers of input variables which are likely correlated with each other, "handcrafted" neurons and connections based on domain knowledge can help mitigate the issue of multi-collinearity in purely data-driven models. Third, in the likely scenarios of imperfect domain knowledge, the model structure can be relaxed to include other similar networks to take advantage of knowledge exploration capabilities of ML models. In this study, the HNN model performance improved significantly after adding DR6. Weights of the edges connected to DR6 were used to infer its physical connotation, thus enabling knowledge discovery. On the flip side, if certain weights are consistently insignificant under various model configurations and data inputs, the corresponding domain knowledge then becomes dubious or ungrounded.

3.4 Potential applications in policy decision support

For municipalities and countries with limited amounts of financial and organizational resources, it is pivotal to identify leverage points to expedite sustainable transitions to a circular economy.[2] This involves the identification and prioritization of feasible, actionable policy levers and technical designs that mobilize social, technical, and financial capital. Such policy levers and decisions can be qualitatively identified from Figure 2 and quantitatively assessed with the HNN model. Specifically, the values of decision variables (DVs) can be manually adjusted to simulate policy decisions, and responses of the HNN model can be treated as quantitative effects of these interventions.

As a quick demonstration, the 11 decision variables (DV1-11) were manually adjusted to increase by 0.5 (constrained to be lower than 1) sequentially, with one at a time, for the U.S., China, and Vietnam to evaluate the corresponding changes in recycling rate. The most effective policy levers turn out to be *Governance Development* and *Processing* as summarized in Table 1, affirming the effectiveness of capacity building in boosting MSW recovery. For the U.S. and China, setting ambitious goals and targets can also help boost recycling rate by 7.7% and 10.1%, respectively. For Vietnam, capacity building and private sector engagement (e.g., extended producer responsibility and public-private partnership) are the most effective decisions, possibly due to a lack of existing infrastructure.[52]

Table 1 The simulated top 3 effective decisions in boosting recycling rates

Country	*Decision 1*	*Decision 2*	*Decision 3*
U.S.	*Goals & Targets* (7.7%)	*Governance Dev.* (5.2%)	*Processing* (1.9%)
China	*Processing* (12.2%)	*Governance Dev.* (10.9%)	*Goals & Targets* (10.1%)
Vietnam	*Pvt. Engagement* (49.4%)	*Governance Dev.* (47.0%)	*Pub. Engagement* (20.0%)

Note: predicted recycling rate increases are presented in parentheses.

More policy insights can be generated by manually adjusting multiple policy decisions at the same time to evaluate their joint effects, such as market development with recycling targets, regulation reinforcement with collection system upgrades, public engagement with economic incentives, etc. These combinations can be easily experimented with in the HNN model.

Another potential application of the HNN model is the analysis of system tipping points, which are certain thresholds of key variables in a system that can lead to directional changes of system behaviors when exceeded. A well-known example in the waste management field is the Kuznets Curve assumption, which assumes a bell-shaped curve of MSW generation as wealth accumulates.[54] To explore similar trends in SWMS, sensitivity analysis can be conducted on key socioeconomic variables by assuming sets of future scenarios and policy trajectories. Once tipping points are identified, their driving forces and propagation throughout the system can be strengthened to accelerate desirable changes or curbed to prevent undesirable system behaviors.

3.5 Limitations and future works

There are a few limitations of the HNN model in its current form. First, the domain knowledge portrayed in Figure 2 needs to be validated and it is naive to use simple sigmoid functions to model the complex nonlinear relationships across diverse SWMS. Second, the dataset is small in size and high in noise level, mainly due to the inherent data availability and quality limitations. As a result, the model training process and model outcomes are relatively sensitive to random data splits (i.e., the training, validation, and testing set splits). Although data limitations lead to imperfect model results, the dataset was purposefully constructed to demonstrate the advantages of the HNN modeling approach. Finally, the model results and policy insights discussed above should be treated as global level observations, rather than regional policy recommendations suited to specific SWMS.

The potential of HNN models in policy support hinges on future improvements that can be made in an iterative manner in both data-driven modeling techniques and science-based domain knowledge. From the data science perspective, additional high-quality observations sourced from diverse socioeconomic backgrounds, especially from developing countries, are indispensable to the generalizability of the model. And feature learning techniques should be implemented to cope with the issue of missing data. Region-specific datasets and costs of policy interventions are also needed to enhance the reliability and practicality of the HNN model recommendations.[53] On the sustainability science side, more rigorous scientific investigations into each node and each edge of Figure 2 are required to specify a robust data generating process and more realistic relationships among the variables, which may require sophisticated probabilistic network models.[55] Moreover, the knowledge discovery capabilities of ML should be leveraged to test different scientific hypotheses to bridge remaining knowledge gaps.

Although the "tunnel" of sustainable waste management is far from being "dug through", this study embarks on a new journey by demonstrating that machine learning with human knowledge is better than machine learning alone at modeling sustainability systems. With an innovative hybrid modeling approach, a collaborative roadmap, and a solid head start on both sides of the "tunnel", we hope this work will inspire many researchers to join this venture.

Supporting Information: decision making context of SWMS, input data and data preparation process, calculation equations of the HNN model, discussions on the algorithm and hybridization treatments, cross-validation for hyperparameter tuning, trained model parameters and training errors, python code with input data files.

Acknowledgement

This research is supported by the Mao Yisheng Fellowship of Carnegie Mellon University to Rui He, and through the CMU-Portugal project "Bee2Waste Crypto" (IDT-COP 45933). The author Ian Scott would like to acknowledge the financial support provided by Fundação para a Ciência e a Tecnologia (FCT) Portugal under the project UIDB/0415s2/2020—Centro de Investigação em Gestão de Informação (MagIC). The authors would like to thank Dr. Scott Matthews for reviewing and editing this paper.

References

(1) Kaza, S.; Yao, L. C.; Bhada-Tata, P.; Van Woerden, F. *What a Waste 2.0 : A Global Snapshot of Solid Waste Management to 2050*; World Bank: Washington, DC., 2018.

(2) Seadon, J. K. Sustainable Waste Management Systems. *J. Clean. Prod.* **2010**, *18* (16–17), 1639–1651. https://doi.org/10.1016/j.jclepro.2010.07.009.

(3) Cobo, S.; Dominguez-Ramos, A.; Irabien, A. From Linear to Circular Integrated Waste Management Systems: A Review of Methodological Approaches. *Resour. Conserv. Recycl.* **2018**, *135* (July 2017), 279–295. https://doi.org/10.1016/j.resconrec.2017.08.003.

(4) Das, S.; Lee, S. H.; Kumar, P.; Kim, K. H.; Lee, S. S.; Bhattacharya, S. S. Solid Waste Management: Scope and the Challenge of Sustainability. *J. Clean. Prod.* **2019**, *228*, 658–678. https://doi.org/10.1016/j.jclepro.2019.04.323.

(5) Chang, N.-B.; Wang, S. F. The Development of an Environmental Decision Support System for Municipal Solid Waste Management. *Comput. Environ. Urban Syst.* **1996**, *20* (3), 201–212. https://doi.org/10.1016/S0198-9715(96)00015-4.

(6) Roberts, K. P.; Turner, D. A.; Coello, J.; Stringfellow, A. M.; Bello, I. A.; Powrie, W.; Watson, G. V. R. SWIMS: A Dynamic Life Cycle-Based Optimisation and Decision Support Tool for Solid Waste Management. *J. Clean. Prod.* **2018**, *196*, 547–563. https://doi.org/10.1016/j.jclepro.2018.05.265.

(7) Sardarmehni, M.; Anchieta, P. H. C.; Levis, J. W. Solid Waste Optimization Life-Cycle Framework in Python (SwolfPy). *J. Ind. Ecol.* **2022**, 1–15. https://doi.org/10.1111/jiec.13236.

(8) Liu, Y.; Sun, C.; Xia, B.; Cui, C.; Coffey, V. Impact of Community Engagement on Public Acceptance towards Waste-to-Energy Incineration Projects: Empirical Evidence from China. *Waste Manag.* **2018**, *76*, 431–442. https://doi.org/10.1016/j.wasman.2018.02.028.

(9) Brooks, A. L.; Wang, S.; Jambeck, J. R. The Chinese Import Ban and Its Impact on Global Plastic Waste Trade. *Sci. Adv.* **2018**, *4* (6), 1–8. https://doi.org/10.1126/sciadv.aat0131.

(10) Jordan, M. I.; Mitchell, T. M. Machine Learning: Trends, Perspectives, and Prospects. *Science (80-.).* **2015**, *349* (6245), 255–260. https://doi.org/10.1126/science.aaa8415.

(11) Abdallah, M.; Abu Talib, M.; Feroz, S.; Nasir, Q.; Abdalla, H.; Mahfood, B. Artificial Intelligence Applications in Solid Waste Management: A Systematic Research Review. *Waste Manag.* **2020**, *109*, 231–246. https://doi.org/10.1016/j.wasman.2020.04.057.

(12) Zhong, S.; Zhang, K.; Bagheri, M.; Burken, J. G.; Gu, A.; Li, B.; Ma, X.; Marrone, B. L.; Ren, Z. J.; Schrier, J.; Shi, W.; Tan, H.; Wang, T.; Wang, X.; Wong, B. M.; Xiao, X.; Yu, X.; Zhu, J.-J.; Zhang, H. Machine Learning: New Ideas and Tools in Environmental Science and Engineering. *Environ. Sci. Technol.* **2021**, *55* (19), acs.est.1c01339. https://doi.org/10.1021/acs.est.1c01339.

(13) Lin, K.; Zhao, Y.; Kuo, J.-H.; Deng, H.; Cui, F.; Zhang, Z.; Zhang, M.; Zhao, C.; Gao, X.; Zhou, T.; Wang, T. Toward Smarter Management and Recovery of Municipal Solid Waste: A Critical Review on Deep Learning Approaches. *J. Clean. Prod.* **2022**, *346* (October 2021), 130943. https://doi.org/10.1016/j.jclepro.2022.130943.

(14) Bakshan, A.; Srour, I.; Chehab, G.; El-Fadel, M.; Karaziwan, J. Behavioral Determinants towards Enhancing Construction Waste Management: A Bayesian Network Analysis.

Resour. Conserv. Recycl. **2017**, *117*, 274–284. https://doi.org/10.1016/j.resconrec.2016.10.006.

(15) Magazzino, C.; Falcone, P. M. Assessing the Relationship among Waste Generation, Wealth, and GHG Emissions in Switzerland: Some Policy Proposals for the Optimization of the Municipal Solid Waste in a Circular Economy Perspective. *J. Clean. Prod.* **2022**, *351* (September 2021), 131555. https://doi.org/10.1016/j.jclepro.2022.131555.

(16) Rafew, S. M.; Rafizul, I. M. Application of System Dynamics Model for Municipal Solid Waste Management in Khulna City of Bangladesh. *Waste Manag.* **2021**, *129*, 1–19. https://doi.org/10.1016/j.wasman.2021.04.059.

(17) Xiao, S.; Dong, H.; Geng, Y.; Tian, X.; Liu, C.; Li, H. Policy Impacts on Municipal Solid Waste Management in Shanghai: A System Dynamics Model Analysis. *J. Clean. Prod.* **2020**, *262*, 121366. https://doi.org/10.1016/j.jclepro.2020.121366.

(18) Pinha, A. C. H.; Sagawa, J. K. A System Dynamics Modelling Approach for Municipal Solid Waste Management and Financial Analysis. *J. Clean. Prod.* **2020**, *269*, 122350. https://doi.org/10.1016/j.jclepro.2020.122350.

(19) Kollikkathara, N.; Feng, H.; Yu, D. A System Dynamic Modeling Approach for Evaluating Municipal Solid Waste Generation, Landfill Capacity and Related Cost Management Issues. *Waste Manag.* **2010**, *30* (11), 2194–2203. https://doi.org/10.1016/j.wasman.2010.05.012.

(20) Karpatne, A.; Atluri, G.; Faghmous, J. H.; Steinbach, M.; Banerjee, A.; Ganguly, A.; Shekhar, ShashiSamatova, N.; Kumar, V. Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data. *IEEE Trans. Knowl. Data Eng.* **2017**, *29* (10), 2318–

2331. https://doi.org/10.1109/tkde.2017.2720168.

(21) Gil, Y. Thoughtful Artificial Intelligence: Forging a New Partnership for Data Science and Scientific Discovery. *Data Sci.* **2017**, *1* (1–2), 119–129. https://doi.org/10.3233/ds-170011.

(22) Rosé, C. P.; McLaughlin, E. A.; Liu, R.; Koedinger, K. R. Explanatory Learner Models: Why Machine Learning (Alone) Is Not the Answer. *Br. J. Educ. Technol.* **2019**, *50* (6), 2943–2958. https://doi.org/10.1111/bjet.12858.

(23) Williams, A.; Kennedy, S.; Philipp, F.; Whiteman, G. Systems Thinking: A Review of Sustainability Management Research. *J. Clean. Prod.* **2017**, *148*, 866–881. https://doi.org/10.1016/j.jclepro.2017.02.002.

(24) Marshall, R. E.; Farahbakhsh, K. Systems Approaches to Integrated Solid Waste Management in Developing Countries. *Waste Manag.* **2013**, *33* (4), 988–1003. https://doi.org/10.1016/j.wasman.2012.12.023.

(25) Mukhtar, E. M.; Williams, I. D.; Shaw, P. J. Visibility of Fundamental Solid Waste Management Factors in Developing Countries. *Detritus* **2018**, *1* (March), 162–173. https://doi.org/10.26403/detritus/2018.16.

(26) Zaman, A. U. Identification of Waste Management Development Drivers and Potential Emerging Waste Treatment Technologies. *Int. J. Environ. Sci. Technol.* **2013**, *10* (3), 455–464. https://doi.org/10.1007/s13762-013-0187-2.

(27) Wilson, D. C. Development Drivers for Waste Management. *Waste Manag. Res.* **2007**, *25* (3), 198–207. https://doi.org/10.1177/0734242X07079149.

(28) Ma, J.; Hipel, K. W. Exploring Social Dimensions of Municipal Solid Waste Management

around the Globe – A Systematic Literature Review. *Waste Manag.* **2016**, *56*, 3–12. https://doi.org/10.1016/j.wasman.2016.06.041.

(29) Meng, X.; Tan, X.; Wang, Y.; Wen, Z.; Tao, Y.; Qian, Y. Investigation on Decision-Making Mechanism of Residents' Household Solid Waste Classification and Recycling Behaviors. *Resour. Conserv. Recycl.* **2019**, *140* (September 2018), 224–234. https://doi.org/10.1016/j.resconrec.2018.09.021.

(30) Babaei, A. A.; Alavi, N.; Goudarzi, G.; Teymouri, P.; Ahmadi, K.; Rafiee, M. Household Recycling Knowledge, Attitudes and Practices towards Solid Waste Management. *Resour. Conserv. Recycl.* **2015**, *102*, 94–100. https://doi.org/10.1016/j.resconrec.2015.06.014.

(31) Laurent, A.; Clavreul, J.; Bernstad, A.; Bakas, I.; Niero, M.; Gentil, E.; Christensen, T. H.; Hauschild, M. Z. Review of LCA Studies of Solid Waste Management Systems - Part II: Methodological Guidance for a Better Practice. *Waste Manag.* **2014**, *34* (3), 589–606. https://doi.org/10.1016/j.wasman.2013.12.004.

(32) Rodrigues, A. P.; Fernandes, M. L.; Rodrigues, M. F. F.; Bortoluzzi, S. C.; Gouvea da Costa, S. E.; Pinheiro de Lima, E. Developing Criteria for Performance Assessment in Municipal Solid Waste Management. *J. Clean. Prod.* **2018**, *186*, 748–757. https://doi.org/10.1016/j.jclepro.2018.03.067.

(33) Turcott Cervantes, D. E.; López Martínez, A.; Cuartas Hernández, M.; Lobo García de Cortázar, A. Using Indicators as a Tool to Evaluate Municipal Solid Waste Management: A Critical Review. *Waste Manag.* **2018**, *80*, 51–63. https://doi.org/10.1016/j.wasman.2018.08.046.

(34) Martinez-Sanchez, V.; Kromann, M. A.; Astrup, T. F. Life Cycle Costing of Waste

Management Systems: Overview, Calculation Principles and Case Studies. *Waste Manag.* **2015**, *36*, 343–355. https://doi.org/10.1016/j.wasman.2014.10.033.

(35) Tchobanoglous, G.; Kreith, F. *Handbook of Solid Waste Management*, Second.; McGraw-Hill, 2002. https://doi.org/10.1036/0071356231.

(36) Wilson, D. *Global Waste Management Outlook*; David C., W., Ed.; United Nations Environment Programme, 2015. https://doi.org/10.18356/765baec0-en.

(37) Govindan, K.; Hasanagic, M. A Systematic Review on Drivers, Barriers, and Practices towards Circular Economy: A Supply Chain Perspective. *Int. J. Prod. Res.* **2018**, *56* (1–2), 278–311. https://doi.org/10.1080/00207543.2017.1402141.

(38) Goodfellow, I.; Bengio, Y.; Courville, A. Chapter 6: Deep Feedforward Networks. In *Deep Learning*; MIT Press, 2016; pp 164–223.

(39) Hornik, K.; Stinchcombe, M.; White, H. Multilayer Feedforward Networks Are Universal Approximators. *Neural Networks* **1989**, *2* (5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8.

(40) Kannangara, M.; Dua, R.; Ahmadi, L.; Bensebaa, F. Modeling and Prediction of Regional Municipal Solid Waste Generation and Diversion in Canada Using Machine Learning Approaches. *Waste Manag.* **2018**, *74*, 3–15. https://doi.org/10.1016/j.wasman.2017.11.057.

(41) Alidoust, P.; Keramati, M.; Hamidian, P.; Amlashi, A. T.; Gharehveran, M. M.; Behnood, A. Prediction of the Shear Modulus of Municipal Solid Waste (MSW): An Application of Machine Learning Techniques. *J. Clean. Prod.* **2021**, *303*, 127053. https://doi.org/10.1016/j.jclepro.2021.127053.

(42) von Rueden, L.; Mayer, S.; Beckh, K.; Georgiev, B.; Giesselbach, S.; Heese, R.; Kirsch, B.; Pfrommer, J.; Pick, A.; Ramamurthy, R.; Walczak, M.; Garcke, J.; Bauckhage, C.; Schuecker, J. Informed Machine Learning -- A Taxonomy and Survey of Integrating Knowledge into Learning Systems. *IEEE Trans. Knowl. Data Eng.* **2019**, 1–20. https://doi.org/10.1109/TKDE.2021.3079836.

(43) Djeumou, F.; Neary, C.; Goubault, E.; Putot, S.; Topcu, U. Neural Networks with Physics-Informed Architectures and Constraints for Dynamical Systems Modeling. **2021**, *168*, 1–15.

(44) Willard, J.; Jia, X.; Xu, S.; Steinbach, M.; Kumar, V. Integrating Scientific Knowledge with Machine Learning for Engineering and Environmental Systems. *ACM Comput. Surv.* **2022**. https://doi.org/10.1145/3514228.

(45) Karniadakis, G. E.; Kevrekidis, I. G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-Informed Machine Learning. *Nat. Rev. Phys.* **2021**, *3* (6), 422–440. https://doi.org/10.1038/s42254-021-00314-5.

(46) He, R.; Sandoval-Reyes, M.; Scott, I.; Semeano, R.; Ferrão, P.; Matthews, S.; Small, M. J. Global Knowledge Base for Municipal Solid Waste Management: Framework Development and Application in Waste Generation Prediction. *J. Clean. Prod.* **2022**, *377* (July), 134501. https://doi.org/10.1016/j.jclepro.2022.134501.

(47) World Bank. World Bank Open Data https://data.worldbank.org/ (accessed Jan 8, 2021).

(48) World Intellectual Property Organization (WIPO). *Global Innovation Index 2021*, 14th ed.; Dutta, S., Lanvin, B., Leon, L. R., Wunsch-Vincent, S., Eds.; WIPO: Geneva, 2021.

(49) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.;

Gimelshein, N.; Antiga, L.; Desmaison, A.; Köpf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; Chintala, S. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Adv. Neural Inf. Process. Syst.* **2019**, *32* (NeurIPS).

(50) Lakhan, C. A Comparison of Single and Multi-Stream Recycling Systems in Ontario, Canada. *Resources* **2015**, *4* (2), 384–397. https://doi.org/10.3390/resources4020384.

(51) Hamilton, W. L.; Ying, R.; Leskovec, J. Representation Learning on Graphs: Methods and Applications. **2017**, 1–24.

(52) Salhofer, S.; Jandric, A.; Soudachanh, S.; Le Xuan, T.; Tran, T. D. Plastic Recycling Practices in Vietnam and Related Hazards for Health and the Environment. *Int. J. Environ. Res. Public Health* **2021**, *18* (8), 4203. https://doi.org/10.3390/ijerph18084203.

(53) Izquierdo-Horna, L.; Kahhat, R.; Vázquez-Rowe, I. Reviewing the Influence of Sociocultural, Environmental and Economic Variables to Forecast Municipal Solid Waste (MSW) Generation. *Sustain. Prod. Consum.* **2022**, *33*, 809–819. https://doi.org/10.1016/j.spc.2022.08.008.

(54) Ercolano, S.; Lucio Gaeta, G. L.; Ghinoi, S.; Silvestri, F. Kuznets Curve in Municipal Solid Waste Production: An Empirical Analysis Based on Municipal-Level Panel Data from the Lombardy Region (Italy). *Ecol. Indic.* **2018**, *93* (December 2017), 397–403. https://doi.org/10.1016/j.ecolind.2018.05.021.

(55) Crane, H.; Dempsey, W. A Statistical Framework for Modern Network Science. *Stat. Sci.* **2021**, *36* (1), 51–67. https://doi.org/10.1214/19-STS759.

Supporting Information for

A novel domain knowledge-informed machine learning approach for modeling solid waste management systems

Rui He[a], Mitchell J. Small[a], Ian Scott[b], Motolani Olarinre[c], Mexitli Sandoval-Reyes[d], Paulo Ferrão[d],*

(*) Correspondence author: iscott@novaims.unl.pt

a Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, United States

b NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Campus de Campolide, 1070-312, Lisbon, Portugal

c Department of Machine Learning, Carnegie Mellon University, Pittsburgh, PA 15213, United States

d IN[+] Center for Innovation, Technology and Policy Research, LARSyS, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal

Nova de Lisboa, Campus de Campolide, 1070-312, Lisbon, Portugal

Number of pages: 32

Number of figures: 7

Number of tables: 4

Contents

1. Methodologies and data

1.1 SWMS decision-making context

Table S1 Variables to model the decision-making context of SWMS

Variable(s)	Code	Description	Assumptions and Notes	References
Demographics	SV1-SV5	Demographic variables that set the demographic background of SWMS	Household size, percentage of population over 65, population density, education level, and urbanization rate are used to describe the demographics	1–3
Culture	SV6-SV7	The lifestyles, moral norms, social influence, and habits of a society	Expense (as % GDP used for households/NPISHs consumption) and energy consumption rate are assumed as proxies to reflect culture	2,4,5
Income	SV8	The purchasing power of households (with GNI per capita as proxy)	Household income sets the socioeconomic background of SWMS	1–3,6
MSW Quantity	SV9	Quantity of MSW generation in a country or municipality	MSW generation sets the technical background of SWMS	2,7,8
Technology	SV10	Technology research and development that improve the efficiency and reduce the cost of technologies	Technology innovation with global innovation index as proxy sets the sociotechnical background of SWMS	9,10
GHGs	SV11	Greenhouse gas emissions (with countries' CO_2 emission as proxy)	GHGs emission sets the environmental background	6,11,12
Pollution	SV12	Toxic and non-toxic impacts to the ecosystems (with PM2.5 air pollution as proxy)	Pollution sets the environmental background of SWMS	6,12–14
GDP	SV13	Economic development (affluence) measured in per capita GDP	GDP sets the economic background of SWMS	2,3,15

MSW Composition	SV14-SV16	Fractions of different components in MSW (e.g., recyclables, compostable, and trash)	MSW composition sets the technical background of SWMS	2,16
Public Attitude	SV17	Public attitude towards waste management and trust in SWMS	Assumed to be correlated with demographics, income, and culture (SV1-SV8)	17–19
Demand	SV18	Demand for waste management (to recover materials and energy)	Higher MSW quantity (due to economies of scale), technology innovation (lower costs), and income are assumed to boost demand	20,21
Public Health	DR1	Public health risk remains a key driver of waste management for developing countries	A main driver for those with high pollution and low GDP	22–24
Public Awareness	DR2	Public awareness of MSW management issues and knowledge that move up the hierarchy of people's priorities as living standards improve	Mainly correlated with pollution, MSW generation, affluence, and public attitude	17,18,22
Environmental Protection	DR3	Phasing out of uncontrolled disposal (in developing countries) and higher environmental standards (in developed countries)	Mainly driven by pollution levels and economic development	22,24
Climate Change	DR4	Leading to carbon reduction policies (e.g., diversion of biodegradable wastes from landfill, energy recovery, and market-based solutions)	Assumed to be correlated with GDP, public awareness, and GHGs emissions	22–24
Market Value	DR5	Value of different collected MSW streams generates business interests and profit motives, providing a livelihood for the urban poor in developing countries	Mainly determined by demand, supply, MSW composition, affluence, and MSW quality (assumed to be correlated with demographic variables)	22,24–26
Goals and Targets	DV1	Paradigm shifts to a circular economy with non-binding goals and targets aiming at higher and better uses of MSW	Assumed to be triggered by climate change, public awareness, and environmental protection drivers	22,27

Legislation	DV2	Legal requirements and binding targets of how MSW streams should be managed and technical standards of treatment technologies	Assumed to be triggered by public health and environmental protection	22,23
Governance Development	DV3	Financial (budget) and institutional capacity building for MSW management	Assumed to be triggered by public health, economic development, public awareness, and environmental protection	22
Market Development	DV4	Business opportunities and market development for secondary materials (e.g., public purchase guidelines favoring recycled materials)	Assumed to be triggered by market value and climate change	20,26
Private Engagement	DV5	Engagement of the private sector, such as extended producer responsibility (EPR) and public-private partnership (PPP)	Assumed to be triggered by market value, climate change, and environmental protection	22,26
Public Engagement	DV6	Engagement of the public sector, such as public education, decision-making engagement, information disclosure, etc.	Assumed to be triggered by public awareness, climate change, and environmental protection	22,26,27
Private Incentives	DV7	Economic incentives to the private sector, including taxes and subsidies (landfill tax, recycling subsidy, carbon credits, etc.)	Assumed to be driven by climate change, but less likely to happen when the market value is high	22,27,28
Public Incentives	DV8	Economic incentives to the SWMS users, such as pay-as-you-throw and save-as-you-throw	Assumed to be driven by market value and climate change, but less prominent with high public awareness	22,27–29
Collection	DV9	More advanced MSW collection system setup (e.g., multi-stream collection systems)	Assumed to be triggered by market value, climate change, and public awareness, and enabled by affluence	8,17,30
Collector	DV10	The inclusion of private sector and informal sector (waste pickers) in waste collection and transportation	Assumed to be driven by value	31,32

Processing	DV11	Processing and storage of MSW as preparation for final treatments (e.g., material recovery facilities, mechanical biological treatments, drop-off, and sorting centers)	Assumed to be driven by market value and climate change, and enabled by affluence	8,32
Goals	CE1	Existence of non-binding goals and targets aiming at higher and better uses of MSW	Results of DV1	22,27
Governance	CE2	Financial (budget) and institutional capacity of MSW management	Assumed to be built by legislation, governance capacity building, private sector engagement, and affluence	22
Public Behavior	CE3	Waste generation, separation, and reduction behaviors of the public	Assumed to be based on public awareness with feedback from various public interfacing policy decisions (DV2, DV6, DV8, and DV9)	1,29
Economic Value	CE4	Value of different processed MSW streams generates business interests and profit motives	Assumed to be market value (DR5) with feedback from policy decisions of DV4, DV5, DV7, DV9, DV10, and DV11	22,24–26
Technical Properties	CE5-CE7	Waste properties that make MSW streams valuable for recovery, including material recoverability (recyclability), nutrients recoverability (biodigestibility), and energy recoverability (WtE potential)	Assumed to be mainly determined by technical decisions (DV9-11) and MSW compositions (SV14-16)	8,33,34
Laws	MO0	Existence of legal requirements and binding targets of how MSW streams should be managed and technical standards of treatment technologies	Results of legislation (DV2)	2,22,23
Final Treatments	MO1-MO5	Major MSW disposal and valorization options (recycle, compost, WtE, landfill, and open dump) as final outcomes of the model	The final treatments are determined based on the CEs	2,8,32

Table S2 Excluded variables of SWMS

Variable Name	Description	Reasons of Exclusion	References
Climate Pattern	Temperature and precipitation patterns that influence waste generation	Slow in change, and non-essential in MSW management decision-making	2
Collection Frequency	The frequency at which MSW streams get collected	Purely a logistical decision, which is out of the scope of this study	8
Contaminants	Contaminant levels in different collected streams that affect the value of MSW streams	Poor data availability, assumed to be related to SV2-SV5	8
Costs	Operation costs, maintenance costs, and capital investment of SWMS	Poor data availability	28,35
Existing Infrastructure	SWMS infrastructure that determines the existing practices of MSW management	Poor data availability	
Geography	Geography of a region that can influence waste collection and transportation	Mainly impacts logistics	
Resource Conservation	Conservation and consumptions of resources as a driver of MSW recovery	Poor data availability	11,12
Revenue	Cost savings and income from waste management	Poor data availability	15,35
Supply	Supply of MSW	Assumed to be the same as MSW Quantity (SV9)	20,21
Trade	International and domestic trades of MSW	Poor data availability	36
Transportation	Transportation and storage of MSW	Mainly impacts logistics	8
Waste Generator	MSW generators, including residential, commercial, and institutional sources	Indifferentiable due to data availability	8

Waste Type	Types of waste according to the EU List of Waste classification	This study only focuses on MSW	37

1.2 Data collection and processing

Data collection and processing were conducted on the foundation of prior research of the authors.[2] Input data, including 16 state variables (SV1-16) and 6 model output variables (MO0-MO5) were collected from the data sources listed in Table S3. As illustrated in the flow chart (Figure S1), 1720 data records were collected from various sources listed in Table S3. Initial data screening was conducted to remove outliers (per capita MSW generation higher than 7 kg/day or lower than 0.01 kg/day) and records that are not MSW generation or MSW disposal. After removal of records with missing data (except for SV14-16, where missing data were replaced with average values), logarithmic transformation was applied to *MSW Quantity*, *Population Density*, *Household Size*, *Energy*, *GDP*, *Income*, *GHGs*, and *Pollution*. Finally, all variables beyond the range of 0 to 1 (SV1, SV3-4, SV7-13) were normalized to the range of 0 to 1 with the equation below.

$$SV'_{ij} = \frac{SV_{ij} - minSV_j}{maxSV_j - minSV_j}$$

where: SV_{ij} is the i^{th} observation of variable SV_j, $minSV_j$ and $maxSV_j$ denote the minimum and maximum values of variable SV_j.

278 records were randomly selected and held out as the testing set. The other records were randomly assigned into training sets (878 records) and the validation sets (220 records) during the 5-fold cross validation (CV) process.

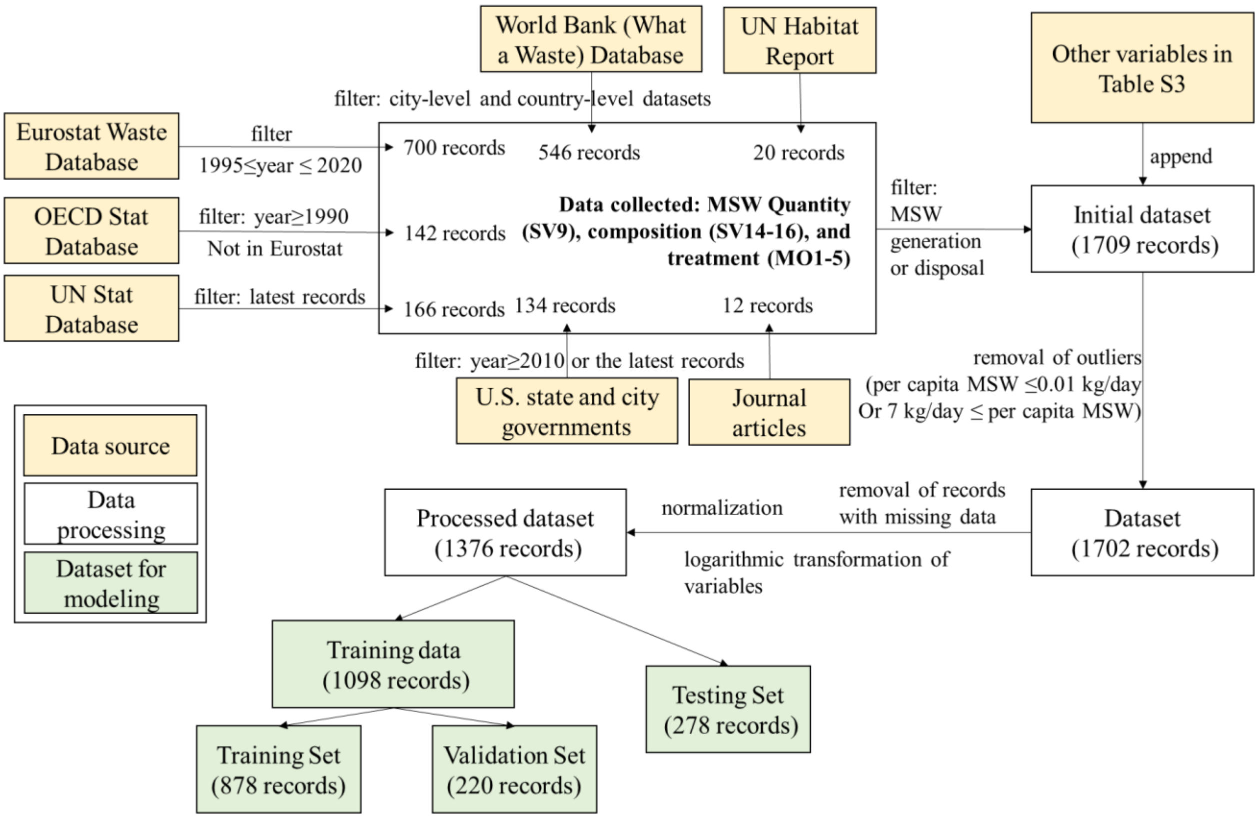


Figure S1 Flow chart of data collection and preparation

Table S3 Collected data for the training of the HNN model

Variable	Description	Unit	Type	Source
Household Size (SV1)	Number of people per household	people/ household	Numerical	UN Data: household size[38]
Age 65+ (SV2)	Percentage of population above 65	%	Numerical	WB Data: Population ages 65+ (% of total population)[6]
Population Density (SV3)	Number of inhabitants per square km	million/km^2	Numerical	Same as SV9, World Bank (WB) Data,[6] and Google
Education (SV4)	Education index of Human Development Index (HDI)		Numerical	UNDP Human Development Index[39]
Urbanization (SV5)	Percentage of people living in urban areas	%	Numerical	WB Data: Urban population (% of total population)[6]
Expense (SV6)	Percentage of GDP used for household and NPISHs consumption	%	Numerical	WB Data: Households and NPISHs final consumption expenditure (% of GDP)[6]
Energy (SV7)	Per capita energy demand of a country	kWh/person	Numerical	UN Data: gross energy demand in million kWh[40]

Income (SV8)	Gross National Income (GNI) per capita	2015 USD /person	Numerical	WB Data: GNI (constant 2015 US$)[6]
MSW Quantity (SV9)	Annual MSW generation	tonnes /year	Numerical	Eurostat Waste Database,[41] OECD Database,[42] UN Stat Database,[43,44] World Bank Database,[45] UN Habitat,[7] U.S. EPA,[46] journal articles,[47–49] and Google search
Technology (SV10)	Technology innovation indexes		Numerical	Global Innovation Index 2021[10]
GHGs (SV11)	CO_2 emissions	Kt	Numerical	WB Data: CO2 emissions (kt)[6]
Pollution (SV12)	PM2.5 air pollution	mg/m^3	Numerical	WB Data: PM2.5 air pollution, mean annual exposure[6]
GDP (SV13)	Per capita GDP of a country	2015 USD /person	Numerical	WB Data: GDP per capita (constant 2015 US$)[6]
Recyclables% (SV14)	Percentage of recyclable waste in MSW	%	Numerical	Same as SV9
Compostable% (SV15)	Percentage of compostable waste in MSW	%	Numerical	Same as SV9
Trash% (SV16)	Other wastes in MSW (1-SV14-SV15)	%	Numerical	Same as SV9
Laws (MO0)	Whether a country has waste regulation or not	Y/N	Binary	UNEP[50]
Recycle (MO1)	Percentage of MSW that is recycled	%	Numerical	Same as SV9
Compost (MO2)	Percentage of MSW that is composted	%	Numerical	Same as SV9
WtE (MO3)	Percentage of MSW that is incinerated or combusted	%	Numerical	Same as SV9
Landfill (MO4)	Percentage of MSW that is landfilled	%	Numerical	Same as SV9
Open Dump (MO5)	Percentage of MSW that is disposed without any control measures	%	Numerical	Same as SV9

1.3 Hybrid neural network (HNN) calculation equations

The calculation equations of the HNN model:

$$SV17 = tanh(SV1 \times \alpha_{1,1} + SV2 \times \alpha_{1,2}^{+} + SV3 \times \alpha_{1,3} + SV4 \times \alpha_{1,4}^{+} + SV5 \times \alpha_{1,5} + SV6 \times \alpha_{1,6} + SV8 \times \alpha_{1,7}^{+} + bias_1)$$

$$SV18 = tanh(SV8 \times \alpha_{2,1}^{+} + SV9 \times \alpha_{2,2} + SV10 \times \alpha_{2,3}^{+} + bias_2)$$

$$DR1 = tanh(SV12 \times \alpha_{3,1}^{+} + SV13 \times \alpha_{3,2}^{-} + bias_3)$$

$$DR2 = tanh(SV9 \times \alpha_{4,1}^{+} + SV12 \times \alpha_{4,2}^{+} + SV13 \times \alpha_{4,3}^{+} + SV17 \times \alpha_{4,4}^{+} + bias_4)$$

$$DR3 = tanh(SV12 \times \alpha_{5,1}^{+} + SV13 \times \alpha_{5,2} + bias_5)$$

$$DR4 = tanh(SV11 \times \alpha_{6,1}^{+} + SV13 \times \alpha_{6,2}^{+} + SV17 \times \alpha_{6,3}^{+} + bias_6)$$

$$DR5 = tanh(SV2 \times \alpha_{7,1}^{+} + SV3 \times \alpha_{7,2}^{+} + SV4 \times \alpha_{7,3}^{+} + SV9 \times \alpha_{7,4} + SV13 \times \alpha_{7,5}^{-} + SV14 \times \alpha_{7,6}^{+} + SV15 \times \alpha_{7,7} + SV16 \times \alpha_{7,8}^{-} + SV18 \times \alpha_{7,9}^{+} + bias_7)$$

$$DR6 = tanh(\sum_{i=1}^{18}\left(SVi \times \alpha_{8,i}\right) + bias_8)$$

$$DV1 = tanh(DR2 \times \alpha_{9,1}^{+} + DR3 \times \alpha_{9,2}^{+} + DR4 \times \alpha_{9,3}^{+} + DR6 \times \alpha_{9,4} + bias_9)$$

$$DV2 = tanh(DR1 \times \alpha_{10,1}^{+} + DR3 \times \alpha_{10,2}^{+} + DR6 \times \alpha_{10,3} + bias_{10})$$

$$DV3 = tanh(DR1 \times \alpha_{11,1}^{+} + DR2 \times \alpha_{11,2}^{+} + DR3 \times \alpha_{11,3}^{+} + DR6 \times \alpha_{11,4} + SV13 \times \alpha_{11,5}^{+} + bias_{11})$$

$$DV4 = tanh(DR4 \times \alpha_{12,1}^{+} + DR5 \times \alpha_{12,2}^{+} + DR6 \times \alpha_{12,3} + bias_{12})$$

$$DV5 = tanh(DR3 \times \alpha_{13,1}^{+} + DR4 \times \alpha_{13,2}^{+} + DR5 \times \alpha_{13,3}^{+} + DR6 \times \alpha_{13,4} + bias_{13})$$

$$DV6 = tanh(DR2 \times \alpha_{14,1}^{+} + DR3 \times \alpha_{14,2}^{+} + DR4 \times \alpha_{14,3}^{+} + DR6 \times \alpha_{14,4} + bias_{14})$$

$$DV7 = tanh(DR4 \times \alpha_{15,1}^{+} + DR5 \times \alpha_{15,2}^{-} + DR6 \times \alpha_{15,3} + bias_{15})$$

$$DV8 = tanh(DR2 \times \alpha_{16,1}^{-} + DR4 \times \alpha_{16,2}^{+} + DR5 \times \alpha_{16,3}^{+} + DR6 \times \alpha_{16,4} + bias_{16})$$

$$DV9 = tanh(DR2 \times \alpha_{17,1}^{+} + DR4 \times \alpha_{17,2}^{+} + DR5 \times \alpha_{17,3}^{+} + DR6 \times \alpha_{17,4} + SV13 \times \alpha_{17,5}^{+} + bias_{17})$$

$$DV10 = tanh(DR5 \times \alpha_{18,1}^{+} + DR6 \times \alpha_{18,2} + SV13 \times \alpha_{18,3}^{-} + bias_{18})$$

$$DV11 = tanh(DR4 \times \alpha_{19,1}^{+} + DR5 \times \alpha_{19,2}^{+} + DR6 \times \alpha_{19,3} + SV13 \times \alpha_{19,4}^{+} + bias_{19})$$

$$CE1 = tanh(DV1 \times \alpha_{20,1}^{+} + bias_{20})$$

$$CE2 = tanh(DV2 \times \alpha_{21,1}^{+} + DV3 \times \alpha_{21,2}^{+} + DV5 \times \alpha_{21,3}^{+} + SV13 \times \alpha_{21,4}^{+} + bias_{21})$$

$$CE3 = tanh(DV2 \times \alpha_{22,1}^{+} + DR2 \times \alpha_{22,2}^{+} + DV6 \times \alpha_{22,3}^{+} + DV8 \times \alpha_{22,4}^{+} + DV9 \times \alpha_{22,5}^{-} + bias_{22})$$

$$CE4 = tanh(DV4 \times \alpha_{23,1}^{+} + DR5 \times \alpha_{23,2}^{+} + DV5 \times \alpha_{23,3}^{+} + DV7 \times \alpha_{23,4}^{+} + DV9 \times \alpha_{23,5}^{+} + DV10 \times \alpha_{23,6}^{+} + DV11 \times \alpha_{23,7}^{+} + bias_{23})$$

$$CE5 = tanh(DV5 \times \alpha_{24,1}^{+} + SV14 \times \alpha_{24,2}^{+} + DV9 \times \alpha_{24,3}^{+} + DV10 \times \alpha_{24,4}^{+} + DV11 \times \alpha_{24,5}^{+} + bias_{24})$$

$$CE6 = tanh(DV9 \times \alpha_{25,1}^{+} + SV15 \times \alpha_{25,2}^{+} + DV10 \times \alpha_{25,3}^{+} + DV11 \times \alpha_{25,4}^{+} + bias_{25})$$

$$CE7 = tanh(DV9 \times \alpha_{26,1}^{+} + DV11 \times \alpha_{26,2}^{+} + SV14 \times \alpha_{26,3} + SV15 \times \alpha_{26,4} + bias_{26})$$

$$MO0 = DV2$$

$$MO1 = logistic(CE1 \times \alpha_{27,1}^{+} + CE2 \times \alpha_{27,2}^{+} + CE3 \times \alpha_{27,3}^{+} + CE4 \times \alpha_{27,4}^{+} + CE5 \times \alpha_{27,5}^{+} + bias_{27})$$

$$MO2 = logistic(CE1 \times \alpha_{28,1}^{+} + CE2 \times \alpha_{28,2}^{+} + CE3 \times \alpha_{28,3}^{+} + CE4 \times \alpha_{28,4} + CE6 \times \alpha_{28,5}^{+} + bias_{28})$$

$$MO3 = logistic(CE1 \times \alpha_{29,1}^{+} + CE2 \times \alpha_{29,2}^{+} + CE3 \times \alpha_{29,3} + CE4 \times \alpha_{29,4} + CE7 \times \alpha_{29,5}^{+} + bias_{29})$$

$$MO4 = logistic(CE1 \times \alpha_{30,1}^{-} + CE2 \times \alpha_{30,2}^{+} + CE3 \times \alpha_{30,3}^{-} + CE4 \times \alpha_{30,4}^{-} + bias_{30})$$

$$MO5 = logistic(CE1 \times \alpha_{31,1}^{-} + CE2 \times \alpha_{31,2}^{-} + CE3 \times \alpha_{31,3}^{-} + CE4 \times \alpha_{31,4}^{-} + bias_{31})$$

where: $\alpha_{i,j}^{-}$ denotes a constrained negative parameter, $\alpha_{i,j}^{+}$ denotes a constrained positive parameter, $\alpha_{i,j}$ denotes an unconstrained parameter, and $bias_j$ denotes a bias term.

1.4 HNN model training algorithm

Training of the HNN model is implemented with the backpropagation algorithm, which is widely used to train all kinds of neural network models including deep learning models.[51] This algorithm is based on the repeated use of the "chain rule" to calculate derivatives and consists of forward computation and backward computation. This algorithm is briefly described below:

1. Forward computation
   a. Write out a model for evaluating the function $\boldsymbol{y} = f(\boldsymbol{x})$, where $\boldsymbol{x}$ is the input (vector), and $\boldsymbol{y}$ is the output (vector)
   b. Define a directed acyclic graph where each variable is a node as the computation graph of the model[51]
   c. Visit each node of the graph in topological order. For each variable u_i with inputs $v_1, v_2, v_3, \dots, v_N$:
      i. (Optional) normalization of inputs
      ii. Compute $u_i = g_i(v_1, v_2, v_3, \dots, v_N)$
      iii. Store the results at the node
2. Backward computation
   a. Initialize $d\boldsymbol{y}/d\boldsymbol{y} = \boldsymbol{1}$
   b. Visit each node of the graph v_j in reverse topological order. Let $u_1, u_2, u_3, \dots, u_M$ denote all the nodes with v_j as an input, Assuming that $\boldsymbol{y} = h(\boldsymbol{u}) = h(u_1, u_2, u_3, \dots, u_M)$, and $u_i = g_i(v_1, v_2, v_3, \dots, v_N)$ for all *i*
      i. We already know $d\boldsymbol{y}/du_i$ for all *i*
      ii. Compute $d\boldsymbol{y}/dv_j = \sum_{i=1}^{M} \frac{dy}{du_i}\frac{du_i}{dv_j}$

3. Return partial derivatives $d\boldsymbol{y}/du_i$ for all variables

With the backbone of backpropagation algorithm above, the HNN model training can be expressed with the pseudo code below:

Initialize model parameters $\boldsymbol{\alpha}$ (including all the $\alpha_{i,j}$ and bias terms)

    For each epoch $e \in \{1, 2,3, \dots, E\}$, do

        For each training data sample $\{\boldsymbol{x}, \boldsymbol{y}\} \in D$, do

            Compute all the nodes and loss function (J) with forward computation

            Compute gradients via backward computation:

            $g_\alpha = \nabla_\alpha J$, where $\nabla_\alpha J$ is the partial derivative $dJ/d\alpha$

        Update parameters: $\alpha = \alpha - r_\alpha g_\alpha$, where r_α is the adaptive learning rate for parameter α

    Evaluate loss function J

Return learnt parameters $\boldsymbol{\alpha}$

1.5 Discussion of the hybridization treatments

In this section, we will discuss the assumptions of the 6 hybridization treatments to the HNN model and why the aforementioned training algorithm is still applicable.

1. **"Hand-crafted" NN structures**

The backpropagation algorithm for calculating gradients is based on repeated use of the "chain rule", which applies to arbitrary computation graphs (as long as they are differentiable). The "hand-crafted" neurons (nodes) and connections (directed edges) do not violate the definition of a computation graph. Thus, the backpropagation algorithm still applies.

However, in a traditional NN model, activation functions imposed on the hidden layers are used to enable simulation of arbitrary nonlinear relationships.[52] When the neurons and connections are "hand-crafted" to convey physical meanings, assumptions have to be made that all the linear relationships and nonlinear transformations approximate the real relationships among the variables (nodes). These assumptions are difficult to make in the absence of domain knowledge for these nodes and connections. Thus, a general nonlinearity assumption is made in this study as a generic starting point. In cases of strong evidence that sigmoid nonlinearity does not hold, other types of equations can be adopted, such as splines or linear relationships.

2. **The weight of parameters can be constrained or fixed based on domain knowledge.**

Directly constraining a weight parameter $\alpha_{i,j}$ to be positive or negative might cause the search space to be non-continuous and cause the model training to be nonconvergent. A better solution is to transform these parameters: $\alpha_{i,j}^{+} = \sqrt{{\alpha_{i,j}}^2}$ (for positive relationships) and $\alpha_{i,j}^{-} = -\sqrt{{\alpha_{i,j}}^2}$ (for negative relationships) to impose science-based restrictions on the model. Exponential transformation is also a viable option, but more attention should be paid to the scale of its gradients since it may dominate the gradients of other (linear) terms.

If a weight parameter $\alpha_{i,j}$ is fixed based on domain knowledge, it is a constant. Then the gradient $g_\alpha = \nabla_\alpha J$ for this parameter is zero. Based on the HNN training algorithm, there will be no update to this parameter.

3. Neurons can be predetermined and fixed based on domain knowledge.

The neurons in the HNN are calculated based on eq. 1 and 2 in the manuscript. However, the form of eq.1 and 2 can be adjusted based on domain knowledge to be any differentiable equations (nonlinear, linear, splines, etc.) as long as the HNN does not violate the definition of a computation graph.

For proven non-monotonic relationships between a node u_i and its input node v_j, transformations of v_j (e.g., polynomial terms) can be added because it is equivalent to inserting new nodes to the computation graph, e.g., $v_{new} = {v_j}^2$. However, if the input node (v_j) is not independent from other inputs (say v_k), interaction terms need to be incorporated as well, e.g., $v_j \cdot v_k$, ${v_j}^2 \cdot v_k$, etc.

In case there is a fixed relationship between a node u_i and its sole input node v_j, the node u_i can be removed from the computation graph to create a simpler model structure of skip-connection neural networks.[53]

As the name suggests, skip-connection NNs allow the outputs of a layer to skip one or more layers in a neural network and feed into the subsequent layers.[53] Existing skip-connection NNs typically feed the entire layer to a subsequent layer (or layers) in deep neural networks such as the ResNet[54] and the DenseNet.[55] However, in this study, the skip-connections are manually added between specific nodes, which is a special case of the skip-connection architecture (equivalent to setting all the other connections at a fixed weight of "0").

4. **Neurons can feed to multiple layers.**

This treatment is exemplified in Figure 2 with the nodes such as *MSW Quantity*, *GDP*, *Recyclables%* feeding to multiple layers. This hybridization treatment can be directly implemented with the skip-connection architecture mentioned above. This treatment can be extended to feeding outputs to the same layer or even previous layers as long as the graph is acyclic.

5. **The neurons with empirical observations in the hidden layers can be added to the loss function, so that the prediction errors on the final outputs and the intermediate neurons can be minimized simultaneously.**

This treatment is exemplified in Figure 2 with the node of *Legislation* (DV2). To include DV2 in the loss function, the computation graph is transformed by adding a skip-connection from DV2 to the output layer. In this case, DV2 can be treated as a calculated variable with a fixed relationship (1-to-1) to the output node *Law* (MO0), which is included in the loss function and trained in the same way as the other output nodes.

6. **Feedback loops of systems can be incorporated by connecting neurons on different layers that share the same physical connotation.**

This feedback loop scenario is exemplified in Figure 2 with the node of *Economic Value* (CE4) and *Market Value* (DR5), which essentially share the same connotation. DR5 triggers certain policy and management decisions, which in return influence the CE4, forming a feedback loop. These loops are prevalent in various systems and can be implemented in HNN models by adding skip-connections.

873

2. Hyperparameter tuning exercise

In this section we provide further details and results on the hyperparameter tuning exercise of the HNN model. We implement 5-fold CV on the training data (1,098 cases) to tune model hyperparameters of optimizer (Figure S2), activation function (Figure S3), and regularization (Figure S4). All these parameters are trained on the training sets (878 cases) and tested on the validation sets (220 cases) in conjunction with different combinations of learning rates (lr) and epochs. A course grid search approach is applied across the ranges of lr (0.1, 0.4, 0.7, 1.0) and results were recorded every 10 epochs up to 250. For each parameter set, the average validation mean squared error (MSE) is reported. The figures below help to identify the best combinations of parameters, which are found at the minimum average validation MSEs.

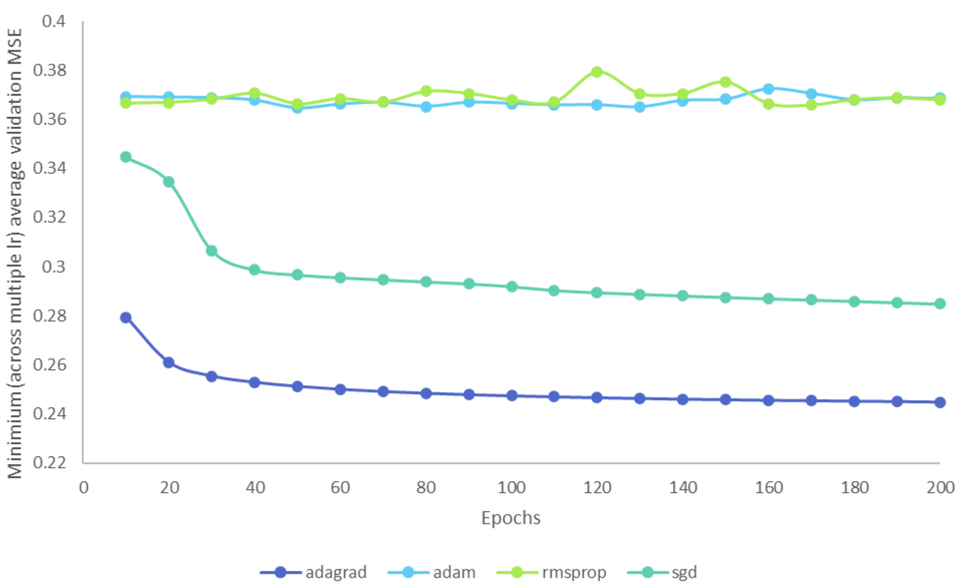


Figure S2 Minimum (across multiple lr) average validation MSEs across 5-fold CV for different combinations of optimizers and epochs, tested against multiple learning rates (0.1, 0.4, 0.7, 1.0)

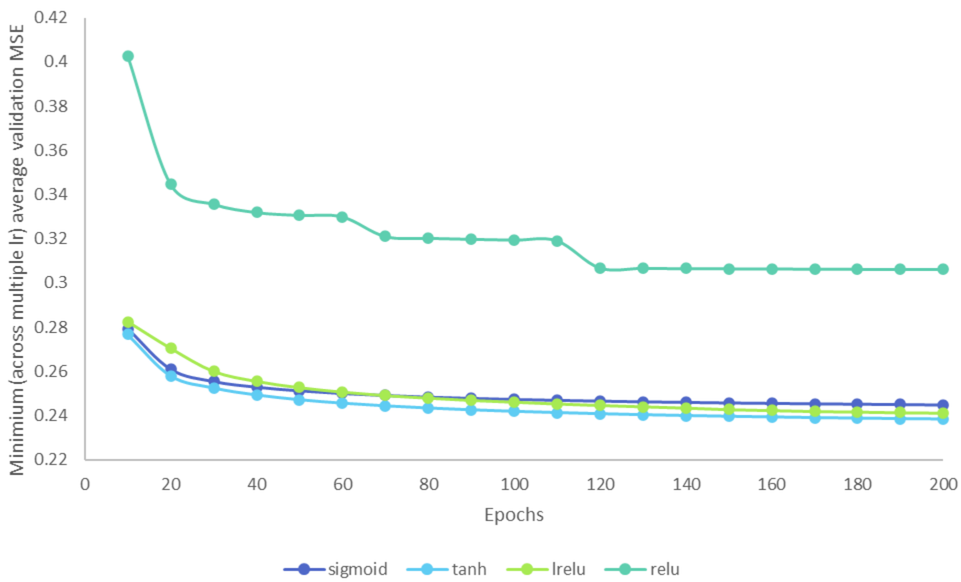
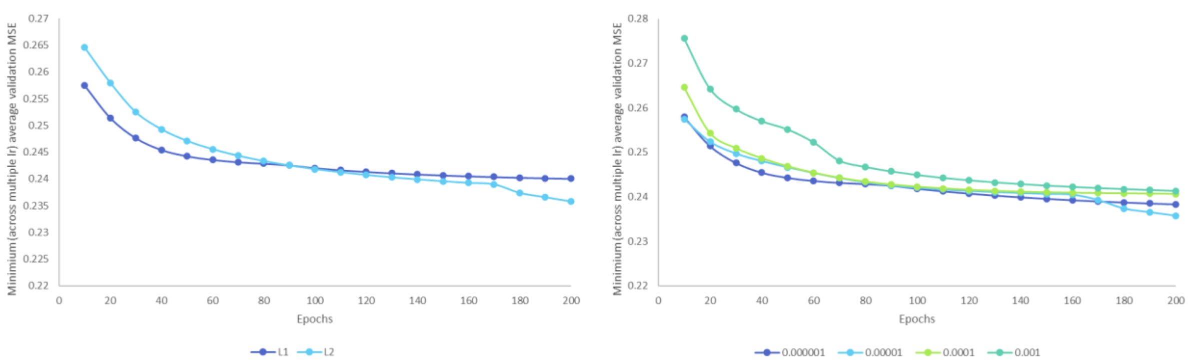


Figure S3 Minimum (across multiple lr) average validation MSEs across 5-fold CV for different combinations of activation functions and epochs, tested against multiple learning rates (0.1, 0.4, 0.7, 1.0)



(a) (b)

Figure S4 Minimum (across multiple lr) average validation MSEs across 5-fold CV for different combinations of a) regularizers and epochs and b) γ and epochs, tested against multiple learning rates (0.1, 0.4, 0.7, 1.0)

After identifying the best parameters based on average validation MSEs, we perform a more fine-tuned grid search across lr (0.1, 0.2, …. 0.8, 1.0) and epochs (10, 20, …., 190, 200) with results presented in Figure S5.

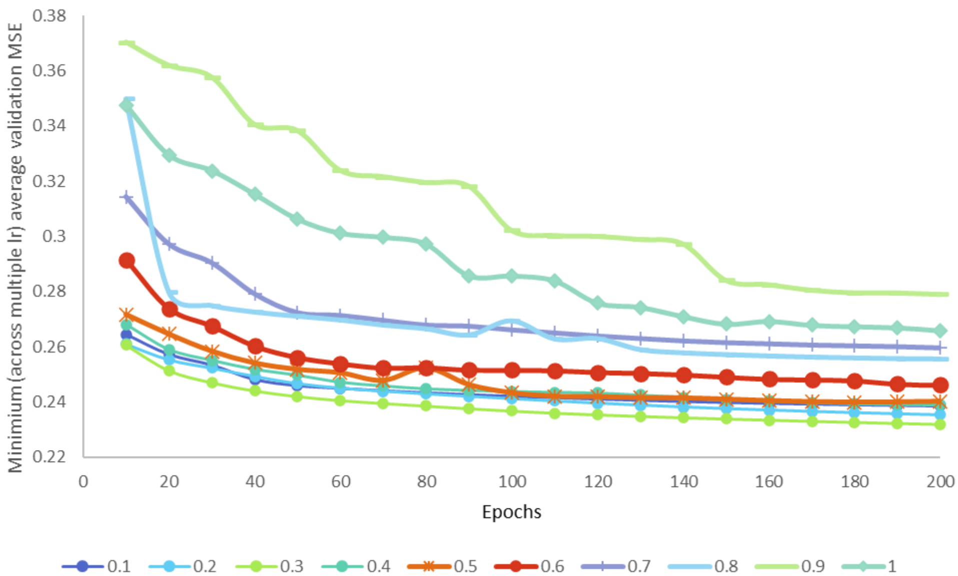


Figure S5 Average validation MSEs across 5-fold CV for different combinations of learning rates and epochs

Overall, the hyperparameter tuning exercise suggests that adagrad, tanh or logistic activation, L2 regularizer (with γ of 0.00001 or 0.000001), learning rate of 0.3, and training time of over 100 epochs could lead to the best HNN model training outcomes. The evolution of training and validation errors is plotted in Figure S6.

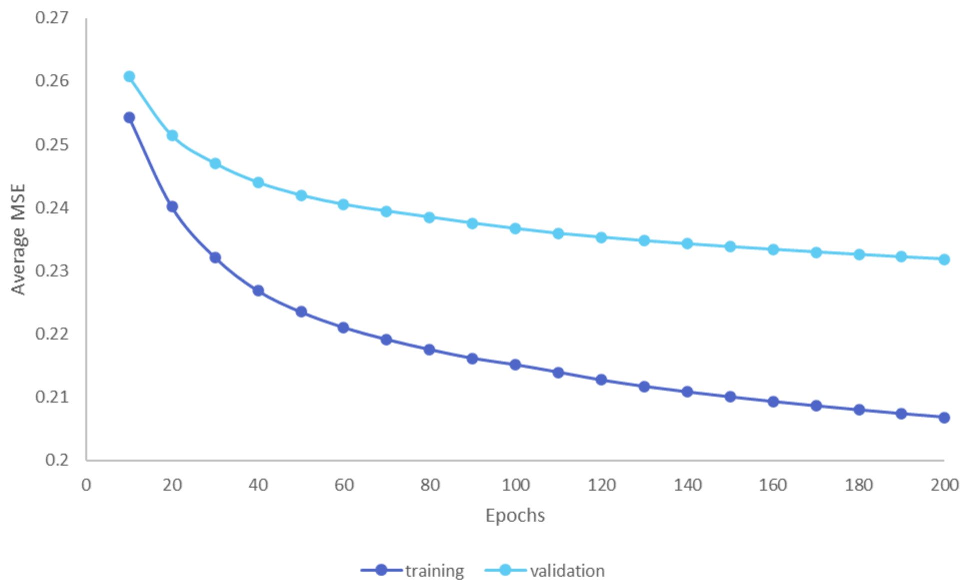


Figure S6 Average training and validation MSEs across 5-fold CV for final parameters

3. Additional results

During the HNN model training stage, there were two sets of parameter weights that led to the minimum testing errors of 0.235 (or below). One set has more insignificant weights (close to "0"), mostly for technical feasibility variables (CE5-7). The other one has fewer insignificant weights (mainly for the *Public Health* driver). The latter one is preferred due to its better interpretability and is presented in Table S4 below.

Table S4 Weights of the trained HNN model parameters

Dependent variable	**Weights of input variables and biases**				
Pub. Attitude	*Household Size*	*Age 65+*	*Pop. Density*	*Education*	*Urbanization*
	-1.823	0.898	0.950	0.347	-0.076
	Expense	*Income*	*Bias*		
	-2.608	0.867	-0.525		
Demand	*Income*	*MSW Quantity*	*Technology*	*Bias*	
	0.475	-0.726	1.354	-0.127	
Pub. Health	*Pollution*	*GDP*	*Bias*		
	0.002	-0.028	0.016		
	MSW Quantity	*Pollution*	*GDP*	*Pub. Attitude*	*Bias*

Pub. Awareness	1.415	0.291	0.335	1.007	-0.824
Env. Protection	*Pollution*	*GDP*	*Bias*		
	2.256	-1.912	-0.385		
Climate Change	*GHGs*	*GDP*	*Pub. Attitude*	*Bias*	
	0.857	1.111	1.171	-1.299	
Market Value	*GDP*	*MSW Quantity*	*Demand*	*Pop. Density*	*Recyclables%*
	-0.577	0.806	0.015	1.785	0.918
	Compostable%	*Trash%*	*Age 65+*	*Education*	*Bias*
	-1.726	-1.772	2.191	0.544	-0.059
Development	*Household Size*	*Age 65+*	*Pop. Density*	*Education*	*Urbanization*
	-2.082	0.049	-0.112	-0.912	-0.495
	Expense	*Energy*	*Income*	*MSW Quantity*	*Technology*
	-1.345	-0.600	-0.887	-0.341	0.803
	GHGs	*Pollution*	*GDP*	*Recyclables%*	*Compostable%*
	-1.389	1.346	0.794	0.047	1.035
	Trash%	*Pub. Attitude*	*Demand*	*Bias*	
	1.122	-1.149	0.379	0.614	
Goals&Target	*Pub. Awareness*	*Env. Protection*	*Climate Change*	*Development*	*Bias*
	1.465	0.010	2.462	1.479	-0.940
Legislation	*Public Health*	*Env. Protection*	*Development*	*Bias*	
	0.001	0.854	-1.736	0.946	
Governance Dev.	*Public Health*	*Pub. Awareness*	*Env. Protection*	*Development*	*GDP*
	0.003	0.413	0.021	1.400	1.031
Market Dev.	*Climate Change*	*Market Value*	*Development*	*Bias*	
	0.433	0.005	-2.057	0.972	
Pvt. Engage.	*Env. Protection*	*Climate Change*	*Market Value*	*Development*	*Bias*
	0.011	1.471	1.461	0.435	1.018
Pub. Engage.	*Pub. Awareness*	*Env. Protection*	*Climate Change*	*Development*	*Bias*
	1.392	4.654	1.166	1.401	-0.891
Pvt. Incentives	*Climate Change*	*Market Value*	*Development*	*Bias*	
	0.231	-1.784	-1.137	0.549	
Pub. Incentives	*Pub. Awareness*	*Climate Change*	*Market Value*	*Development*	*Bias*
	-2.753	0.011	0.023	0.676	-0.594
Collection	*Pub. Awareness*	*Climate Change*	*Market Value*	*Development*	*GDP*
	1.186	1.114	0.861	-0.292	1.730
Collector	*Market Value*	*Development*	*GDP*	*Bias*	
	0.408	-0.472	-0.220	-0.531	
Processing	*Climate Change*	*Market Value*	*Development*	*GDP*	*Bias*
	2.990	0.381	0.870	0.536	-0.053
Goals	*Goals&Targets*	*Bias*			
	1.875	0.249			
Governance	*Legislation*	*Governance Dev.*	*Pvt. Engage.*	*GDP*	*Bias*
	0.012	0.956	0.997	2.072	-0.776
Pub. Behavior	*Legislation*	*Pub. Awareness*	*Pub. Engage.*	*Pub. Incentives*	*Collection*

	0.090	0.034	1.671	0.831	-1.567
Econ. Value	*Market Dev.*	*Pvt. Engage.*	*Pvt. Incentives*	*Market Value*	*Collection*
	0.720	0.108	0.927	0.000	0.013
	Collector	*Processing*	*Bias*		
	0.018	0.876	0.833		
Recyclability	*Pvt. Engage.*	*Recyclables%*	*Collection*	*Collector*	*Processing*
	1.263	1.039	1.142	0.669	2.100
Biodigestibility	*Collection*	*Compostable%*	*Collector*	*Processing*	*Bias*
	0.726	1.840	0.451	0.240	-1.285
WtE Potential	*Collection*	*Processing*	*Recyclables%*	*Compostable%*	*Bias*
	0.788	1.376	-0.748	0.952	-0.415
Recycle	*Goals*	*Governance*	*Pub. Behaviors*	*Value*	*Recyclability*
	0.237	0.872	0.695	0.301	0.797
Compost	*Goals*	*Governance*	*Pub. Behaviors*	*Value*	*Biodigestibility*
	0.284	1.581	2.166	0.318	1.431
WtE	*Goals*	*Governance*	*Pub. Behaviors*	*Value*	*WtE Potential*
	0.836	0.908	0.398	-0.503	1.250
Landfill	*Goals*	*Governance*	*Pub. Behaviors*	*Value*	*Bias*
	-1.621	0.274	-1.604	-2.201	-0.950
Open Dump	*Goals*	*Governance*	*Pub. Behaviors*	*Value*	*Bias*
	-0.554	-2.421	-0.009	-0.036	-1.171

Note: the bias terms for *Governance Dev.* (-0.680), *Collection* (-0.635), *Pub. Behavior* (-0.308), *Recyclability* (-0.432), *Recycle* (-2.269), *Compost* (-2.183), and *WtE* (-1.814).

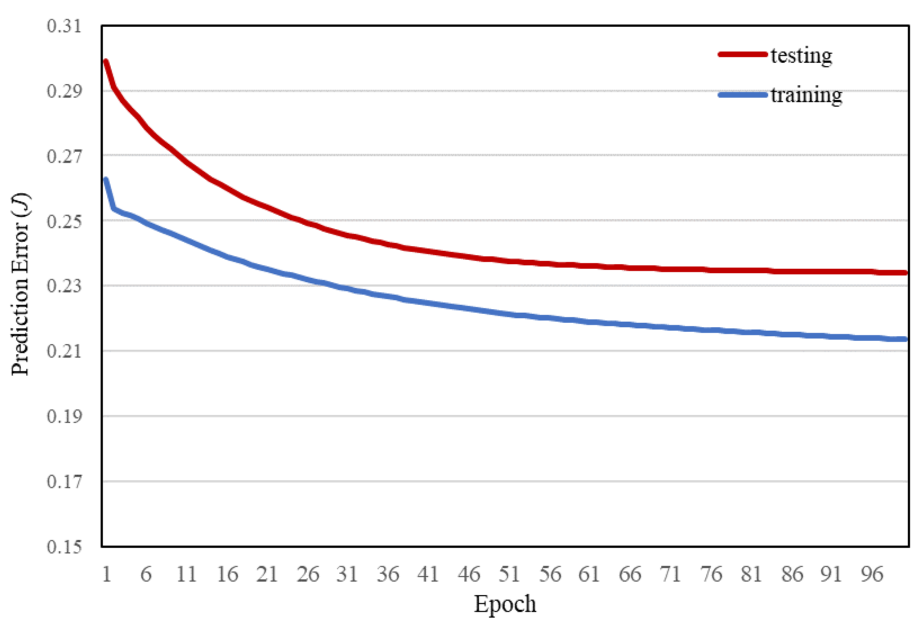


Figure S7 The HNN model training and testing errors (MSEs) over 100 epochs

922

4. References

(1) Ma, J.; Hipel, K. W. Exploring Social Dimensions of Municipal Solid Waste Management around the Globe – A Systematic Literature Review. *Waste Manag.* **2016**, *56*, 3–12. https://doi.org/10.1016/j.wasman.2016.06.041.

(2) He, R.; Sandoval-Reyes, M.; Scott, I.; Semeano, R.; Ferrão, P.; Matthews, S.; Small, M. J. Global Knowledge Base for Municipal Solid Waste Management: Framework Development and Application in Waste Generation Prediction. *J. Clean. Prod.* **2022**, *377* (July), 134501. https://doi.org/10.1016/j.jclepro.2022.134501.

(3) Izquierdo-Horna, L.; Kahhat, R.; Vázquez-Rowe, I. Reviewing the Influence of Sociocultural, Environmental and Economic Variables to Forecast Municipal Solid Waste (MSW) Generation. *Sustain. Prod. Consum.* **2022**, *33*, 809–819. https://doi.org/10.1016/j.spc.2022.08.008.

(4) Kaplan Mintz, K.; Henn, L.; Park, J.; Kurman, J. What Predicts Household Waste Management Behaviors? Culture and Type of Behavior as Moderators. *Resour. Conserv. Recycl.* **2019**, *145* (January), 11–18. https://doi.org/10.1016/j.resconrec.2019.01.045.

(5) Botetzagias, I.; Dima, A. F.; Malesios, C. Extending the Theory of Planned Behavior in the Context of Recycling: The Role of Moral Norms and of Demographic Predictors. *Resour. Conserv. Recycl.* **2015**, *95*, 58–67. https://doi.org/10.1016/j.resconrec.2014.12.004.

(6) World Bank. World Bank Open Data https://data.worldbank.org/ (accessed Jan 8, 2021).

(7) UN-Habitat. *Solid Waste Management in the World's Cities*; Earthscan: London, 2010; Vol. 26.

(8) Roberts, K. P.; Turner, D. A.; Coello, J.; Stringfellow, A. M.; Bello, I. A.; Powrie, W.; Watson, G. V. R. SWIMS: A Dynamic Life Cycle-Based Optimisation and Decision Support Tool for Solid Waste Management. *J. Clean. Prod.* **2018**, *196*, 547–563. https://doi.org/10.1016/j.jclepro.2018.05.265.

(9) Agamuthu, P.; Khidzir, K. M.; Hamid, F. S. Drivers of Sustainable Waste Management in Asia. *Waste Manag. Res.* **2009**, *27* (7), 625–633. https://doi.org/10.1177/0734242X09103191.

(10) World Intellectual Property Organization (WIPO). *Global Innovation Index 2021*, 14th ed.; Dutta, S., Lanvin, B., Leon, L. R., Wunsch-Vincent, S., Eds.; WIPO: Geneva, 2021.

(11) Rodrigues, A. P.; Fernandes, M. L.; Rodrigues, M. F. F.; Bortoluzzi, S. C.; Gouvea da Costa, S. E.; Pinheiro de Lima, E. Developing Criteria for Performance Assessment in Municipal Solid Waste Management. *J. Clean. Prod.* **2018**, *186*, 748–757. https://doi.org/10.1016/j.jclepro.2018.03.067.

(12) Turcott Cervantes, D. E.; López Martínez, A.; Cuartas Hernández, M.; Lobo García de Cortázar, A. Using Indicators as a Tool to Evaluate Municipal Solid Waste Management: A Critical Review. *Waste Manag.* **2018**, *80*, 51–63. https://doi.org/10.1016/j.wasman.2018.08.046.

(13) Laurent, A.; Clavreul, J.; Bernstad, A.; Bakas, I.; Niero, M.; Gentil, E.; Christensen, T. H.; Hauschild, M. Z. Review of LCA Studies of Solid Waste Management Systems - Part II: Methodological Guidance for a Better Practice. *Waste Manag.* **2014**, *34* (3), 589–606. https://doi.org/10.1016/j.wasman.2013.12.004.

(14) Ripa, M.; Fiorentino, G.; Vacca, V.; Ulgiati, S. The Relevance of Site-Specific Data in

Life Cycle Assessment (LCA). The Case of the Municipal Solid Waste Management in the Metropolitan City of Naples (Italy). *J. Clean. Prod.* **2017**, *142*, 445–460. https://doi.org/10.1016/j.jclepro.2016.09.149.

(15) Zaman, A. U. Identification of Waste Management Development Drivers and Potential Emerging Waste Treatment Technologies. *Int. J. Environ. Sci. Technol.* **2013**, *10* (3), 455–464. https://doi.org/10.1007/s13762-013-0187-2.

(16) Karak, T.; Bhagat, R. M.; Bhattacharyya, P. Municipal Solid Waste Generation, Composition, and Management: The World Scenario. *Crit. Rev. Environ. Sci. Technol.* **2012**, *42* (15), 1509–1630. https://doi.org/10.1080/10643389.2011.569871.

(17) Meng, X.; Tan, X.; Wang, Y.; Wen, Z.; Tao, Y.; Qian, Y. Investigation on Decision-Making Mechanism of Residents' Household Solid Waste Classification and Recycling Behaviors. *Resour. Conserv. Recycl.* **2019**, *140* (September 2018), 224–234. https://doi.org/10.1016/j.resconrec.2018.09.021.

(18) Babaei, A. A.; Alavi, N.; Goudarzi, G.; Teymouri, P.; Ahmadi, K.; Rafiee, M. Household Recycling Knowledge, Attitudes and Practices towards Solid Waste Management. *Resour. Conserv. Recycl.* **2015**, *102*, 94–100. https://doi.org/10.1016/j.resconrec.2015.06.014.

(19) Lakhan, C. Exploring the Relationship between Municipal Promotion and Education Investments and Recycling Rate Performance in Ontario, Canada. *Resour. Conserv. Recycl.* **2014**, *92*, 222–229. https://doi.org/10.1016/j.resconrec.2014.07.006.

(20) Matter, A.; Ahsan, M.; Marbach, M.; Zurbrügg, C. Impacts of Policy and Market Incentives for Solid Waste Recycling in Dhaka, Bangladesh. *Waste Manag.* **2015**, *39*, 321–328. https://doi.org/10.1016/j.wasman.2015.01.032.

(21) Das, S.; Lee, S. H.; Kumar, P.; Kim, K. H.; Lee, S. S.; Bhattacharya, S. S. Solid Waste Management: Scope and the Challenge of Sustainability. *J. Clean. Prod.* **2019**, *228*, 658–678. https://doi.org/10.1016/j.jclepro.2019.04.323.

(22) Wilson, D. C. Development Drivers for Waste Management. *Waste Manag. Res.* **2007**, *25* (3), 198–207. https://doi.org/10.1177/0734242X07079149.

(23) Govindan, K.; Hasanagic, M. A Systematic Review on Drivers, Barriers, and Practices towards Circular Economy: A Supply Chain Perspective. *Int. J. Prod. Res.* **2018**, *56* (1–2), 278–311. https://doi.org/10.1080/00207543.2017.1402141.

(24) Marshall, R. E.; Farahbakhsh, K. Systems Approaches to Integrated Solid Waste Management in Developing Countries. *Waste Manag.* **2013**, *33* (4), 988–1003. https://doi.org/10.1016/j.wasman.2012.12.023.

(25) Gobbi, C. Designing the Reverse Supply Chain: The Impact of the Product Residual Value. *Int. J. Phys. Distrib. Logist. Manag.* **2011**, *41* (8), 768–796. https://doi.org/10.1108/09600031111166429.

(26) Magazzino, C.; Falcone, P. M. Assessing the Relationship among Waste Generation, Wealth, and GHG Emissions in Switzerland: Some Policy Proposals for the Optimization of the Municipal Solid Waste in a Circular Economy Perspective. *J. Clean. Prod.* **2022**, *351* (September 2021), 131555. https://doi.org/10.1016/j.jclepro.2022.131555.

(27) Seadon, J. K. Sustainable Waste Management Systems. *J. Clean. Prod.* **2010**, *18* (16–17), 1639–1651. https://doi.org/10.1016/j.jclepro.2010.07.009.

(28) Martinez-Sanchez, V.; Kromann, M. A.; Astrup, T. F. Life Cycle Costing of Waste Management Systems: Overview, Calculation Principles and Case Studies. *Waste Manag.*

2015, *36*, 343–355. https://doi.org/10.1016/j.wasman.2014.10.033.

(29) Vorobeva, D.; Scott, I. J.; Oliveira, T.; Neto, M. Adoption of New Household Waste Management Technologies: The Role of Financial Incentives and pro-Environmental Behavior. *J. Clean. Prod.* **2022**, *362* (May), 132328. https://doi.org/10.1016/j.jclepro.2022.132328.

(30) Bing, X.; Bloemhof, J. M.; Ramos, T. R. P.; Barbosa-Povoa, A. P.; Wong, C. Y.; van der Vorst, J. G. A. J. Research Challenges in Municipal Solid Waste Logistics Management. *Waste Manag.* **2016**, *48*, 584–592. https://doi.org/10.1016/j.wasman.2015.11.025.

(31) Kaza, S.; Yao, L. C.; Bhada-Tata, P.; Van Woerden, F. *What a Waste 2.0 : A Global Snapshot of Solid Waste Management to 2050*; World Bank: Washington, DC., 2018.

(32) Wilson, D. *Global Waste Management Outlook*; David C., W., Ed.; United Nations Environment Programme, 2015. https://doi.org/10.18356/765baec0-en.

(33) Zhang, J.; Qin, Q.; Li, G.; Tseng, C. H. Sustainable Municipal Waste Management Strategies through Life Cycle Assessment Method: A Review. *J. Environ. Manage.* **2021**, *287* (November 2020), 112238. https://doi.org/10.1016/j.jenvman.2021.112238.

(34) Götze, R.; Boldrin, A.; Scheutz, C.; Astrup, T. F. Physico-Chemical Characterisation of Material Fractions in Household Waste: Overview of Data in Literature. *Waste Manag.* **2016**, *49*, 3–14. https://doi.org/10.1016/j.wasman.2016.01.008.

(35) Pinha, A. C. H.; Sagawa, J. K. A System Dynamics Modelling Approach for Municipal Solid Waste Management and Financial Analysis. *J. Clean. Prod.* **2020**, *269*, 122350. https://doi.org/10.1016/j.jclepro.2020.122350.

(36) Brooks, A. L.; Wang, S.; Jambeck, J. R. The Chinese Import Ban and Its Impact on Global

Plastic Waste Trade. *Sci. Adv.* **2018**, *4* (6), 1–8. https://doi.org/10.1126/sciadv.aat0131.

(37) Eurostat. *Guidance on Classification of Waste According to EWC-Stat Categories*; 2010.

(38) United Nations. Household Size and Composition https://www.un.org/development/desa/pd/data/household-size-and-composition (accessed Jan 8, 2021).

(39) UNDP. Human Development Index https://hdr.undp.org/data-center/human-development-index#/indicies/HDI.

(40) United Nations. UN Data http://data.un.org/Default.aspx (accessed Jan 8, 2021).

(41) Eurostat. Eurostat Waste Database https://ec.europa.eu/eurostat/web/waste/data/database (accessed Jan 8, 2021).

(42) OECD. Municipal waste https://data.oecd.org/waste/municipal-waste.htm (accessed Sep 1, 2021). https://doi.org/10.1787/89d5679a-en.

(43) United Nations Statistics Division. Total amount of municipal waste collected http://data.un.org/Data.aspx?q=municipal+waste&d=ENV&f=variableID%3A1814 (accessed Sep 1, 2021).

(44) United Nations Statistics Division. Municipal waste treatment in selected cities https://unstats.un.org/unsd/envstats/Questionnaires/2020/Tables/Municipal waste treatment at city level in selected cities.xlsx (accessed Sep 1, 2021).

(45) World Bank. What a Waste Global Database https://datacatalog.worldbank.org/search/dataset/0039597 (accessed Sep 1, 2021).

(46) US EPA. U.S. State and Local Waste and Materials Characterization Reports

https://www.epa.gov/facts-and-figures-about-materials-waste-and-recycling/us-state-and-local-waste-and-materials (accessed Aug 1, 2021).

(47) Wang, H.; Wang, C. Municipal Solid Waste Management in Beijing: Characteristics and Challenges. *Waste Manag. Res.* **2013**, *31* (1), 67–72. https://doi.org/10.1177/0734242X12468199.

(48) Denafas, G.; Ruzgas, T.; Martuzevičius, D.; Shmarin, S.; Hoffmann, M.; Mykhaylenko, V.; Ogorodnik, S.; Romanov, M.; Neguliaeva, E.; Chusov, A.; Turkadze, T.; Bochoidze, I.; Ludwig, C. Seasonal Variation of Municipal Solid Waste Generation and Composition in Four East European Cities. *Resour. Conserv. Recycl.* **2014**, *89*, 22–30. https://doi.org/10.1016/j.resconrec.2014.06.001.

(49) Ogwueleka, T. C. Survey of Household Waste Composition and Quantities in Abuja, Nigeria. *Resour. Conserv. Recycl.* **2013**, *77*, 52–60. https://doi.org/10.1016/j.resconrec.2013.05.011.

(50) Kumar, S.; Ugirashebuja, E.; Carnwath, L.; Tamminen, T.; Boyd, D. *Environmental Rule of Law - First Global Report*; United Nations Environment Programme, 2019.

(51) Goodfellow, I.; Bengio, Y.; Courville, A. Chapter 6: Deep Feedforward Networks. In *Deep Learning*; MIT Press, 2016; pp 164–223.

(52) Hornik, K.; Stinchcombe, M.; White, H. Multilayer Feedforward Networks Are Universal Approximators. *Neural Networks* **1989**, *2* (5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8.

(53) Yamanaka, J.; Kuwashima, S.; Kurita, T. Fast and Accurate Image Super Resolution by Deep CNN with Skip Connection and Network in Network. In *Proceedings of the*

International Conference on Neural Information Processing; Guangzhou, China, 2017; Vol. 10635 LNCS, pp 217–225. https://doi.org/10.1007/978-3-319-70096-0_23.

(54) He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* **2015**, *2016-Decem*, 770–778. https://doi.org/10.1109/CVPR.2016.90.

(55) Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K. Q. Densely Connected Convolutional Networks. *Am. J. Vet. Res.* **2016**, *39* (9), 1442–1446.