A hierarchical Bayesian regression model that reduces uncertainty in material demand predictions

DOI: 10.1111/jiec.13339

METHODS ARTICLE

JOURNAL OF INDUSTRIAL ECOLOGY WILEY

A hierarchical Bayesian regression model that reduces uncertainty in material demand predictions

Karan Bhuwalka[1] | Eunseo Choi[2] | Elizabeth A. Moore[1] | Richard Roth[1] | Randolph E. Kirchain[1] | Elsa A. Olivetti[3]

[1]Materials Research Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

[2]Institute for Data, Systems and Society, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

[3]Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Correspondence
Karan Bhuwalka, 77 Massachusetts Ave. E19-695, Cambridge, MA 02139, USA.
Email: bhuwalka@mit.edu

Editor Managing Review: Gang Liu

Funding information
This project received no external funding.

Abstract

Predictions of metal consumption are vital for criticality assessments and sustainability analyses. Although demand for a material varies strongly by region and end-use sector, statistical models of demand typically predict demand using regression analyses at an aggregated global level ("fully pooled models"). "Un-pooled" regression models that predict demand at a disaggregated country or regional level face challenges due to limited data availability and large uncertainty. In this paper, we propose a Bayesian hierarchical model that can simultaneously identify heterogeneous demand parameters (like price and income elasticities) for individual regions and sectors, as well as global parameters. We demonstrate the model's value by estimating income and price elasticity of copper demand in five sectors (Transportation, Electrical, Construction, Manufacturing, and Other) and five regions (North America, Europe, Japan, China, and Rest of World). To validate the benefits of the Bayesian approach, we compare the model to both a "fully pooled" and an "un-pooled" model. The Bayesian model can predict global demand with similar uncertainty as a fully pooled regression model, while additionally capturing regional heterogeneity in income elasticity of demand. Compared to un-pooled models that predict demand for individual countries and sectors separately, our model reduces the uncertainty of parameter estimates by more than 50%. The hierarchical Bayesian modeling approach we propose can be used for various commodities, improving material demand projections used to study the impact of policies on mining sector emissions and informing investment in critical material production.

KEYWORDS
Bayesian models, copper, demand, industrial ecology, metals

1 | INTRODUCTION

The world is increasingly dependent on elements from the periodic table once thought rare or trace, driven by the need for sustainable technologies (Klinger, 2020; Vidal et al., 2018). Firms and governments alike are trying to improve the characterization and projection of materials production and demand. These projections are used to understand resource stress (Elshkaki et al., 2016; Graedel et al., 2015; Nassar et al., 2012) and the environmental implications of material consumption (Van der Voet et al., 2019). Unexpected high demand for materials can increase costs, which can influence companies' decisions to make capital investments, commit to circularity, or expand their businesses (Kuipers et al., 2018; Leader et al., 2019; Månberger & Stenqvist, 2018). Demand forecasts are largely conducted at an aggregated global level even though there is significant heterogeneity in demand dynamics at various regions and sectors (He & Small, 2022). Moreover, demand forecasts typically have large uncertainty within and across studies (Watari et al., 2021). For demand estimates to be effective in guiding decision-making, we need novel forecast methods that reduce uncertainty while accounting for regional and sectoral differences.

Models of future material demand are typically informed by two approaches: (1) stock dynamic modeling and (2) inflow-driven modeling. Stock dynamic models (or "bottom up models") calculate material demand as a derivative of the changing stocks of various applications that contain the material. On the other hand, inflow-driven models (or "top down models") find econometric relationships between overall material demand in a region (typically, globally) and input variables such as population and income.

Stock dynamic approaches are used to estimate demand for metals (Deetman et al., 2018; Gerst, 2009; Glöser et al., 2013; Watari et al., 2020) usually in specific smaller-scale case studies like that for South Africa (Kapur & Graedel, 2006), Switzerland (Bader et al., 2011), China (Dong et al., 2019; Zhang et al., 2015), and the United States (He & Small, 2022; Wang et al., 2022). A comprehensive stock dynamics model of the world material stocks requires sufficient data about how much material is used in most major applications, in all regions. Therefore, although stock dynamics models are more detailed and capture heterogeneity between regions and sectors, they are limited by data constraints and a large number of required assumptions (Schipper et al., 2018).

Due to these limitations, the most common approach for estimating demand involves inflow-driven modeling (Müller et al., 2014; Watari et al., 2021). Of the 63 demand projection studies reviewed by Watari et al. (2021), 66% used an inflow-driven approach, while only 26% applied a stock dynamics approach. These models estimate material consumption in each year as a function of variables such as population, income, and price (Elshkaki et al., 2018; Schipper et al., 2018). Inflow-driven approaches have the advantage of accessible aggregate data at the national or global level (Liu et al., 2013) as well as transparency due to the fewer required assumptions. Within the category of inflow-driven models, regression models are commonly used by researchers to analyze the demand for key materials, such as copper (Ciacci et al., 2020; Dong et al. 2019.; Fisher et al., 1972; Kuipers et al., 2018; Van der Voet et al., 2019), aluminum (Elshkaki et al., 2020), and steel (Dhar et al., 2020). Most regression models involve estimating key parameters like income elasticity (i.e., percentage change in material demand resulting from a 1% change in economic growth) and price elasticity (i.e., percentage change in demand resulting from a 1% price changes) from historic data (Crompton, 2015; Fernandez, 2018; Pei & Tilton, 1999). This relationship of demand with income and price is then used to project future demand under various scenarios using scenario analyses (Börjeson et al., 2006).

A major limitation of regression models is that they lack appropriate inflow data in many cases. As a result, they are less useful for detailed regional analysis (Schipper et al., 2018). It is important to understand demand across locations because demand dynamics can differ significantly between developed and developing countries (Ayres et al., 2003; Krausmann et al., 2009), which can be lost in aggregated global demand values (He & Small, 2022). Due to data limitations, regression models that estimate demand across individual countries find large uncertainty in estimates such as price elasticity of demand. For instance, only 23% of the price elasticity values estimated by Fernandez (2018) were statistically significant. The key issue in regression modeling for material demand, therefore, is the trade-off resulting from choosing the level of data pooling. If demand is analyzed at a localized level (i.e., data from individual regions and sectors are treated separately or "un-pooled"), there is larger uncertainty in estimated parameters because of a smaller amount of available inflow data and a greater variation within countries. Ifdata are aggregated and demand is analyzed at a global level (i.e., the demand data are aggregated or "pooled") to reduce uncertainty, heterogeneous trends across different countries and sectors are lost.

In this paper, we apply a Bayesian hierarchical modeling approach to overcome the aforementioned challenges with choosing the level of data pooling in regression modeling of demand (Gelman et al., 2013). Bayesian regression models permit the use of smaller datasets through hierarchical modeling (Vehtari et al., 2017) which allows sharing of information across heterogeneous sites. Hierarchical models estimate local parameters (i.e., price elasticity in a specific country) as well as global parameters (i.e., aggregate price elasticity). The global parameter value (which is estimated by pooling data across sites) acts as a default (or prior) for the local parameter values and therefore reduces uncertainty in local parameter estimates. Therefore, Bayesian hierarchical models make predictions with lower uncertainty and greater accuracy than classical least squares regression, which is itself a special case of hierarchical models (Gelman, 2006).

Bayesian modeling has already been adopted in the literature to predict demand under uncertainty for other systems such as electricity (Wang et al., 2017) and water (Zhang et al., 2019). Bayesian approaches have also been used to reduce uncertainty in material flow analyses when little data are available (Dong et al, 2022; Lupton & Allwood, 2018). In these instances, Bayesian techniques allowed a more informed decision-making by reducing uncertainty while accounting for spatial variation. However, to the best of our knowledge, Bayesian hierarchical regression modeling has not yet been applied to analyze material demand.

The Bayesian hierarchical model we introduce estimates the price and income elasticity of demand (i.e., demand for a material is modeled as a function of material price and GDP). To demonstrate the benefit of our model in cases with low inflow data, we apply the model on copper demand data for five regions and five sectors from 2000 to 2014 (only 15 data points per region-sector combination). We compare our model against an "un-pooled" model where demand parameters for each region–sector combination are estimated independently, as well as against a "fully pooled" model that estimates demand parameters at a global level. While we demonstrate how our approach can improve inflow-based regression models, we cannot compare directly against stock dynamics models which are used in cases with different modeling goals and data availability.

We find significant reduction in the uncertainty of the price elasticities (2.3× lower) and income elasticities (1.6× lower) when using a hierarchical model over an un-pooled model. As prediction uncertainty typically compounds over time, this leads to a very large reduction in uncertainty (>10×) in demand prediction over a 25-year time horizon. Finally, when compared to a fully pooled regression model that only captures global values of income elasticities, a hierarchical model can capture heterogeneity in different regions' response to GDP growth without any increase in uncertainty.

2 | METHODS

We model copper demand in five regions and five sectors as a function of price and income (real gross domestic product or GDP). Our goal is to demonstrate how a Bayesian hierarchical demand model helps reduce uncertainty in material demand estimates compared to linear regression models. We compare the estimated price and income elasticities of demand for each region–sector combination from three candidate models: (i) the Bayesian hierarchical model we propose in this paper, (ii) an "un-pooled" model where the demand in each region and sector is modeled independently, and (iii) a "fully pooled" model where demand is modeled at a global level. In this section, we describe the variables included in our model and how the model form compares with the literature. We then examine the three candidate models and the differences between them. Finally, we describe the data we use to train our models.

2.1 | Model variables

As commonly found in other demand analyses, we model demand as a log-linear relationship between demand, price, and income. We make small modifications to the Pei and Tilton's "simple" metal demand that specifies demand for a particular metal (D_t) in a period as a function of its own price (P_t), the price of a substitute (SP_t,) and income measured by GDP (GDP_t) (Pei & Tilton, 1999). The variable choices and model transformations we make are discussed in detail in Supporting Information Section A.

Specifically, we model the change in log-demand in any region–sector combination as a linear function of the change in the log-price and the change in the region's log-GDP in that period:

$$\Delta \ln\left(D_{s_i,r_j,t}\right) = \beta^{o}_{s_i,r_j} + \beta_{s_i,r_j} \Delta \ln\left(P'_t\right) + \beta^{\text{GDP}}_{s_i,r_j} \Delta \ln\left(\text{GDP}_{r_j,t}\right) \tag{1}$$

The subscript s_i is the sector index for sector i, r_j the region index for region j and year t. $D_{s_i,r_j,t}$, the total copper demand in year t for each of the 25 region–sector combinations.

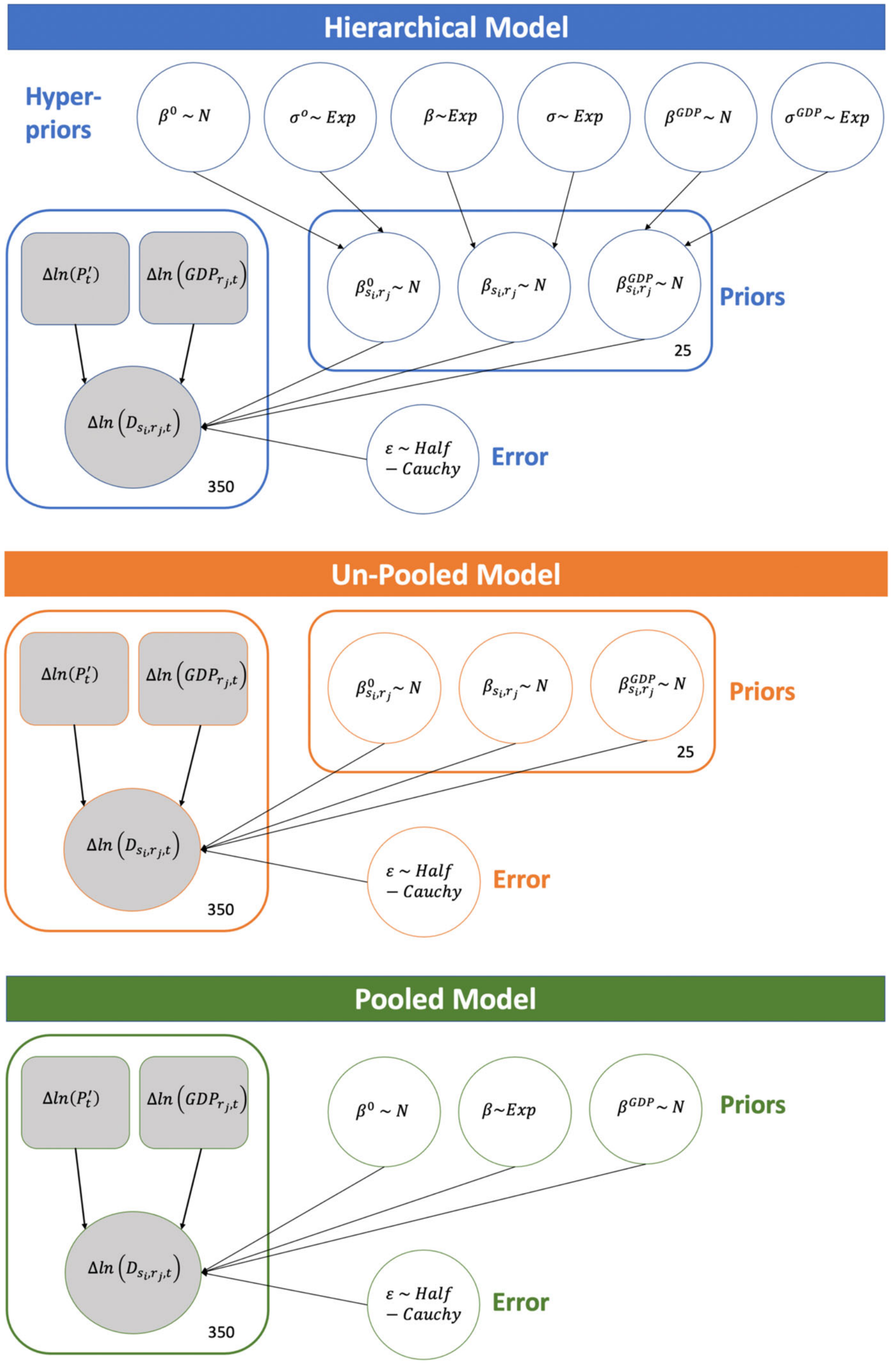
The intercept, $\beta^{o}_{s_i,r_j}$, measures the temporal effect, that is, change in demand due to technology growth or change in consumer preferences. It specifically captures the change in the demand growth rate for a material in the absence of price or income changes. For example, while technology growth and efficiency in manufacturing reduces material demand over time, change in consumer preferences can increase or decrease demand.

The terms β_{s_i,r_j} are the sector- and region-specific price elasticities. The price used here, $P'_t = \frac{P_t + 2*P_{t-1} + P_{t-2}}{4}$, is the trailing 3-year average copper cathode price (similar to Pei and Tilton). We do not use metal price at year t directly because it takes time for manufacturers to cut material intensity in response to price and the demand is more likely impacted by long-term price. Since an increase in raw material price should cause manufacturers to use alternative materials or dematerialize, the price elasticity β_{s_i,r_j} is expected to be negative.

Finally, $\beta^{\text{GDP}}_{s_i,r_j}$ is the income elasticity; it measures the change in the demand due to a percentage change in real GDP. Typically, as income level increases, copper use either increases or stays constant. The income elasticity $\beta^{\text{GDP}}_{s_i,r_j}$ is therefore expected to be non-negative.

2.1.1 | Bayesian hierarchical model

A Bayesian hierarchical model allows for partial pooling of information, providing the option to pool data from different local groups (regions, sectors, or both in our case). Stated differently, multilevel or hierarchical models can separately estimate the predictive effects of an individual predictor



215

216

217

FIGURE 1 Graphical representation of the Bayesian hierarchical model (top), the un-pooled model (middle), and the pooled model (bottom). The grey circle represent the independent (predicted) variable in the model, that is, $\Delta \ln(D_{s_i,r_j,t})$, the grey squares represent the dependent (predictor) variables, that is, $\Delta \ln(P_t')$ and $\Delta \ln(GDP_{r_j,t})$. The white circles represent model parameters, including priors, hyperpriors, and the error term. N() refers to the normal distribution and Exp() refers to the exponential distribution.

and its group-level mean. For this reason, hierarchical models are contextual; they can account for effect specific to each local group, but also the global mean, which further informs the local estimate.

Figure 1 shows a graphical representation of the Bayesian hierarchical model with the dependent and independent variables, as well as parameters (the "priors" and "hyperpriors"). Priors are probability distributions that reflect existing knowledge or belief about uncertain value of a parameter. Priors add value in that it allows for a Bayesian model to estimate the credible interval for a parameter faster without having to search over the entire real number space. Priors inform the "first guess" of the parameter values, essentially improving parameter estimation when data are limited. In our model, we use wide probability distributions as priors to minimize bias in the parameter estimates. (See Supporting Information

TABLE 1 Priors and hyperpriors used in the Bayesian hierarchical model

Priors (local)	Hyperpriors (global)
$\beta^{o}_{s_i,r_j} \sim$ Normal (β^0, σ^o)	$\beta^0 \sim$ Normal $(0,\ 1)$ $\sigma^o \sim$ Exp (1)
$\beta_{s_i,r_j} \sim$ Normal (β, σ)	$\beta \sim -$Exp (1) $\sigma \sim$ Exp (1)
$\beta^{GDP}_{s_i,r_j} \sim$ Normal $(\beta^{GDP}, \sigma^{GDP})$	$\beta^{GDP} \sim$ Normal $(0,\ 1)$ $\sigma^{GDP} \sim$ Exp (1)

Section F to see model sensitivity to change in these priors. Given that we have small data, we assessed the influence of these priors to our model results.)

Hyperpriors are used in our Bayesian hierarchical model to initialize the model such that it finds better mean and standard deviations for each prior (Gelman, 2006). The hyperpriors are informed by the literature but are chosen such that they are non-informative (have wide probability density range). We use a $-Exp\,(1)$ prior for the hyperprior for the global price elasticity β. The 95% confidence interval of an $-Exp\,(1)$ distribution between -3 and 0. As we expect the price elasticity to always be negative, the hyperprior has no support in the positive real region. Moreover, the value of zero has a maximum probability in an exponential distribution which ensures we do not bias the price elasticity estimates. Typical values of global price elasticity fall comfortably within the -3 to 0 range defined by our hyperpriors. We use a normal hyperprior for income elasticity and intercept that is centered around zero so that the posterior estimates are not biased by the prior we assign. Pei and Tilton find the average income elasticity across regions to be 0.066, which comfortably falls within the 95% confidence range of the hyperprior β^{GDP} between -2 and 2 (Pei & Tilton, 1999).

To estimate the parameters, we apply the No-U-Turn (NUTs) algorithm with PyMC3, a gradient-based extension to Hamiltonian Monte Carlo Sampling algorithm that improves efficiency of Markov chain Monte Carlo (MCMC) methods. Models in our experiments are run on two chains. We run 3000 iterations in addition to 1000 discarded samples that were used for model tuning. More details are in Supporting Information Section E. The model likelihood function $p(y_{s_i,r_j,t}|\theta)$ is fit to the data assuming normally distributed errors, that is,

$$p(y_{s_i,r_j,t}|\theta) = N\left(\beta^{o}_{s_i,r_j} + \beta_{s_i,r_j}\Delta\ln\left(P'_t\right) + \beta^{GDP}_{s_i,r_j}\Delta\ln\left(\mathrm{GDP}_{r_j,t}\right), \varepsilon\right) \tag{2}$$

where

$y_{s_i,r_j,t} = \Delta\ln\left(D_{s_i,r_j,t}\right)$
θ is the parameter vector $[\beta^{o}_{s_i,r_j}, \beta_{s_i,r_j}, \beta^{GDP}_{s_i,r_j}, \beta^0, \sigma^o, \beta, \sigma, \beta^{GDP}, \sigma^{GDP}, \varepsilon]\forall i, j$ containing the priors ($\beta^{o}_{s_i,r_j}, \beta_{s_i,r_j}, \beta^{GDP}_{s_i,r_j}$) and hyperpriors ($\beta^0, \sigma^o, \beta, \sigma, \beta^{GDP}, \sigma^{GDP}$). See Table 1 for more details about priors.
s_i = specific sector, $i \in$ {Transportation, Electrical, Construction, Manufacturing and Other}
r_j = specific region, $j \in$ {China, North America, Europe, Japan, Rest of the World}
t = year
ε is the variance of the model error. It is initialized with a prior $\varepsilon \sim Half - Cauchy(5)$

The joint posterior probability $P(\theta|\, y_{s_i,r_j,t})$ for the hierarchical model can be written as follows:

$$P\left(\theta|\, y_{s_i,r_j,t}\right) \propto \prod_{i=1}^{N_r}\prod_{j=1}^{N_s}\prod_{t=1}^{T}\left\{p\left(y_{s_i,r_j,t}|\theta\right)\right\} \times \left\{p\left(\beta^{o}_{s_i,r_j}|\beta^0, \sigma^o\right) p\left(\beta_{s_i,r_j}|\beta, \sigma\right) p\left(\beta^{GDP}_{s_i,r_j}|\beta^{GDP}, \sigma^{GDP}\right)\right\} \times p\left(\beta^0\right) p\left(\sigma^o\right) p\left(\beta\right) p\left(\sigma\right) p\left(\beta^{GDP}\right) p\left(\sigma^{GDP}\right) \tag{3}$$

where

$N_r = 5$ = number of regions considered
$N_s = 5$ = number of sectors considered
$T = 14$ = numbers of years of annual data

and $p(y_{s_i,r_j,t}|\theta)$ is the data likelihood that the model fits (same as Equation 2 above) The rest of the terms in the equation that are multiple with the likelihood are the prior distributions (posterior = likelihood × prior).

Section E in the Supporting Information validates the estimation of the posterior was done correctly by using the posterior-predictive checks outlined by Gabry et al. (2019). In Supporting Information Section E we compare the posterior predicted demand with the actual data and also plot the model error. We find that the error has an average value close to 0.

2.1.2 | Alternative models

In this paper, we compare the Bayesian hierarchical modeling approach outlined above with a "fully pooled" and "un-pooled" approach described in the following sections.

Un-pooled model

If we run individual regression models to estimate demand for each region–sector combination, we are running an "un-pooled" model:

$$\Delta \ln\left(D_{s_i,r_j,t}\right) = \beta^{o}_{s_i,r_j} + \beta_{s_i,r_j}\Delta \ln\left(P'_t\right) + \beta^{\mathrm{GDP}}_{s_i,r_j}\Delta \ln\left(\mathrm{GDP}_{r_j,t}\right) \tag{4}$$

An un-pooled model does not sample global parameters, which can lead to overfitting.

They typically have more uncertainty because we are not sharing information across regions or sectors or their combinations. Pei and Tilton use individual regression models to analyze demand in 18 countries and can be considered to fall under our definition of an un-pooled model (Pei & Tilton, 1999).

A classical un-pooled regression model has uniform completely uninformative priors for the individual parameters ($\beta^{o}_{s_i,r_j}, \beta_{s_i,r_j}, \beta^{\mathrm{GDP}}_{s_i,r_j} \sim U(-\infty,\infty)$). To best simulate this, we initialize parameters with centered (zero mean), uninformative priors (initialized with a very large standard deviation). The main difference between the two models is that the un-pooled model does not estimate a group-level mean (see Figure 1; does not estimate $\beta, \beta^0, \beta^{\mathrm{GDP}}, \sigma, \sigma^{\mathrm{GDP}}, \sigma^0$). Since there is no group-level mean, each site's estimate for parameter values is completely independent of the other sites' parameter estimates.

$$\beta^{o}_{s_i,r_j} \sim \text{Normal}\ (0,\ 100)$$

$$\beta_{s_i,r_j} \sim \text{Normal}\ (0,\ 100)$$

$$\beta^{\mathrm{GDP}}_{s_i,r_j} \sim \text{Normal}\ (0,\ 100)$$

The joint posterior for the un-pooled model can be written as follows:

$$\mathrm{P}\left(\theta | y_{r,s,t}\right) \propto \prod_{i=1}^{Nr}\prod_{j=1}^{Ns}\prod_{t=1}^{T}\left\{p\left(y_{r,s,t}|\beta^{o}_{s_i,r_j}, \beta_{s_i,r_j}, \beta^{\mathrm{GDP}}_{s_i,r_j}\right)\right\} p\left(\beta^{o}_{s_i,r_j}\right) p\left(\beta_{s_i,r_j}\right) p\left(\beta^{\mathrm{GDP}}_{s_i,r_j}\right) \tag{5}$$

Fully pooled model

A regression run on global demand data with no region-specific effects is "fully pooled" because it combines all the local information into one regression. Mathematically, a fully pooled model only samples global parameters, ignoring difference nuances across local groups in the data (see Figure 1; does not estimate $\beta^{o}_{s_i,r_j}, \beta_{s_i,r_j}, \beta^{\mathrm{GDP}}_{s_i,r_j}$). By only estimating global parameters, these models reduce uncertainty but lose site-specific heterogeneous information. For the fully pooled model, we assign the same priors for β, β^0, and β^{GDP} as in the case of the hierarchical model.

$$\Delta \ln\left(D_{s_i,r_j,t}\right) = \beta^{0} + \beta\Delta \ln\left(P'_t\right) + \beta^{\mathrm{GDP}}\Delta \ln\left(\mathrm{GDP}_{r_j,t}\right) \tag{6}$$

where

$$\beta^{0} \sim \text{Normal}\ (0,\ 1)$$

$$\beta \sim -\text{Exp}\ (1)$$

$$\beta^{\mathrm{GDP}} \sim \text{Normal}\ (0,\ 1)$$

For both un-pooled and fully pooled models, we use the same data and the same set-up for estimation (same hardware, sampling runs, etc.).

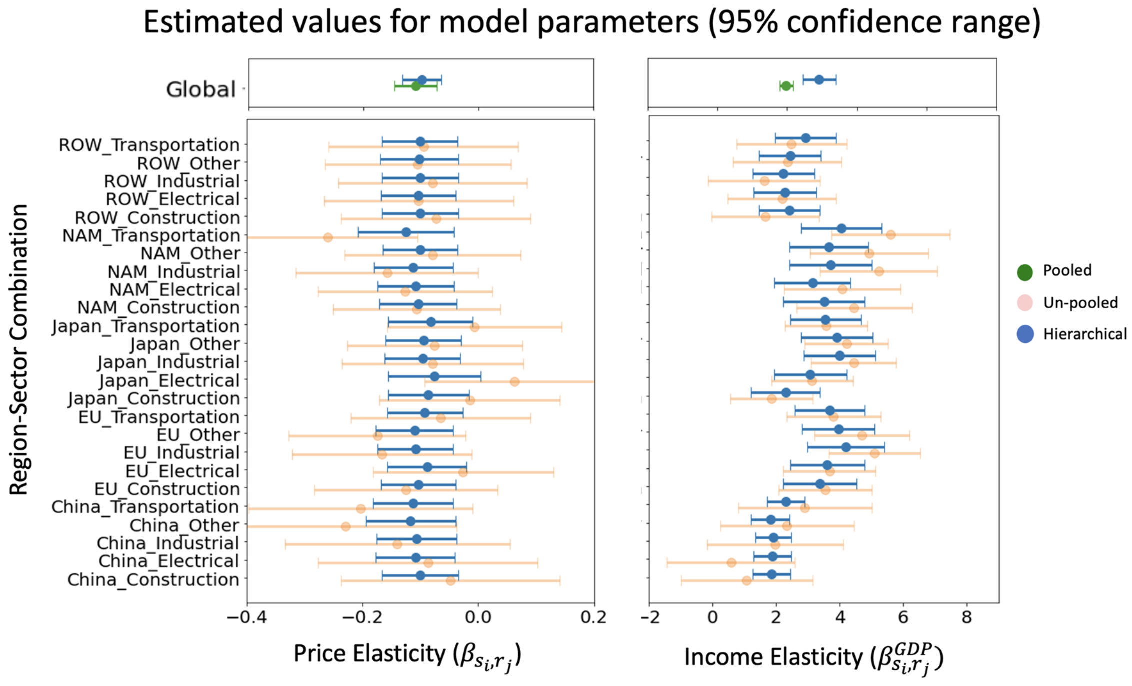


FIGURE 2 Local parameter estimates for un-pooled (orange) and hierarchical (blue) model are displayed along the *x*-axis, with the hierarchical model values overlayed on the un-pooled model values. Circle represents parameter mean and the bar represents the width of 95% confidence interval for the parameter. The region–sector combination is displayed on the *y*-axis. The global estimates for the pooled (green) and hierarchical (blue) model are also displayed. The hierarchical model estimates both local and global parameters, the pooled model estimates only global parameters, while the un-pooled model estimates only local parameters. The two figures correspond to price and income elasticity respectively. The values for the intercept are in Supporting Information Figure S1. The 95% confidence intervals for price and income elasticities from the un-pooled model are wider than the hierarchical model, that is, there is reduced uncertainty in the estimates from the hierarchical model. The values used to generate this figure are present in Supporting Information Table S1.

2.2 | Data

The data we use in our models consist of annual copper demand from 2000 to 2014 broken down by five sectors (Transportation, Electrical, Construction, Manufacturing, and Other) and five regions (North America, Europe, Japan, China, and Rest of World). The demand data measures refined consumption of copper including direct use of scrap. We use historic copper price data (S&P Global Market Intelligence, 2016) and real GDP at constant 2010 USD (United Nations, 2010). The demand data we use consist of 25 region–sector combinations (5 regions and 5 sectors), each of which has only 15 years of annual data. In total, we have 375 data points for demand (15 years × 25 region–sector combinations). We estimate 81 parameters from this data (6 global parameters + 25 region–sector combinations × 3 covariates: intercept, price, and GDP). The disaggregated demand data are from proprietary sources, but the aggregated global demand data are presented in Supporting Information Section C. All other data used (on price and GDP) are also presented in the Supporting Information Section C.

2.2.1 | Results

In this section, we compare the values of price elasticity and income elasticity between the hierarchical and un-pooled model, specifically focusing on the difference in uncertainty of the estimates. The differences in the intercept values are discussed in Supporting Information Section B. We also compare the hierarchical model with a fully pooled model, to evaluate the predictive cost of the additional site-specific parameter information we gain in a hierarchical model.

2.2.2 | Price elasticity

In the hierarchical model, all region–sector combinations have very similar negative price elasticities β_{s_i,r_j} (Figure 2; Supporting Information Table S1). The mean values for all region–sector combination is −0.104 and has a 95% confidence range from −0.076 to −0.13. On average, a 1%

increase in price leads to a 0.104% decrease in demand. The fact that all local price elasticities have similar estimates means that there was very little heterogeneity in price elasticity across region–sector combinations.

Such a low negative price elasticity is consistent with the findings in the literature. Evans and Lewis calculated a long-run copper price elasticity between −0.0909 and −0.1177 (Evans & Lewis, 2005). Other estimates for global copper long-term price elasticity are −0.2 from Fisher et al. (1972) and −0.44 by Dallas Fed (Stuermer, 2017). Pei and Tilton calculate price elasticity for 18 countries and find only 4 that are non-zero (at a 95% confidence level) with an average value of 0.066 (Pei & Tilton, 1999). None of these analyses considered sectoral disaggregation of demand.

The un-pooled model has much larger variation in price elasticity estimates than the hierarchical model. We define the uncertainty in price elasticity for each model (both un-pooled and hierarchical) as the average of the standard deviation of each β_{s_i,r_j}:

$$U_p = \frac{\sum_{i=1}^{5} \sum_{j=1}^{5} \sigma\left(\beta_{s_i,r_j}\right)}{25} \tag{8}$$

The value of U_p for un-pooled model is 0.081, compared to a U_p of 0.035 for the hierarchical model. The uncertainty for the hierarchical model is 2.3× lesser than that for the un-pooled model.

When comparing the hierarchical model and fully pooled model (Figure 2; Supporting Information Table S2), we see no major difference in estimated price elasticities. The global price elasticity estimate β in the hierarchical model has a mean value −0.104 and standard deviation (SD) of 0.018. This estimate is similar to the price elasticity estimated by the fully pooled model (Mean = −0.091; SD = 0.018). Despite additionally estimating 25 site-specific price effects, the hierarchical model still lends a global price elasticity estimate with a similar standard deviation as the fully pooled model.

Income elasticity

As opposed to what we observed for the price elasticities, we find significant heterogeneity between sites for income elasticity values ($\beta^{GDP}_{s_i,r_j}$) estimated by hierarchical model (Figure 2; Supporting Information Table S1). We find the income elasticity to be much stronger than the price elasticity (mean values for $\beta^{GDP}_{s_i,r_j}$range from 1.82 to 4.22, which are more than 10× larger in magnthan estimates for $-\beta_{s_i,r_j}$). This means that the demand for copper is driven by economic development more than by price considerations. Crucially, developing countries (China, RoW) have a smaller value for income elasticity than developed countries (NAM, Japan, EU). The mean value of $\beta^{GDP}_{s_i,r_j}$ for $\forall i,\ j \in \{\text{China, RoW}\}$ ranges from 1.82 to 3.00. Comparatively, the mean value of $\beta^{GDP}_{s_i,r_j}$ for $\forall i,\ j \in \{\text{EU, Japan, NAM}\}$ ranges from 2.30 to 4.22. This result suggests that with an equivalent percentage increase in GDP, developed countries have a larger percentage increase in demand for copper. This is likely because the product mix used by consumers in developed countries require a larger amount of copper than those in developing countries. For example, developed countries consume more electricity, buy more cars, and consequently use more copper for wiring. There is evidence in the literature that supports this finding that developed countries have a stronger demand response to increase in income levels than developing countries. Pei and Tilton find the average income elasticity of copper in developed countries to be 1.450 and in developing countries to be −0.008 (Pei & Tilton, 1999).

For the income elasticity as well, the hierarchical model leads to a significant reduction in uncertainty over the un-pooled model. We define the uncertainty in income elasticity for each model (un-pooled and hierarchical) as:

$$U_{GDP} = \frac{\sum_{i=1}^{5} \sum_{j=1}^{5} \sigma\left(\beta^{GDP}_{s_i,r_j}\right)}{25} \tag{9}$$

The value of U_{GDP} for the un-pooled model is 0.843, compared to a U_{GDP} value of 0.513 for the hierarchical model. Therefore, the uncertainty is reduced by 1.6× due to hierarchical pooling. Noticeably, the uncertainty reduction is lower for the income elasticity than the price elasticities, due to the larger underlying heterogeneity for income elasticity (Meager, 2019).

When comparing the hierarchical model to the fully pooled model (Figure 2; Supporting Information Table S2), we find that the global value of income elasticity (β^{GDP}has a Mean = 2.91 and SD = 0.23) is more uncertain than the income elasticity in a fully pooled model (Mean = 1.96; SD = .093). The fully pooled model has lower deviations in parameter estimates but is unable to capture the heterogeneity between sites that we observe from the local estimates for $\beta^{GDP}_{s_i,r_j}$ in the hierarchical model. A fully pooled model would estimate the same income elasticity for all regions, that is, an increase in GDP at a global level would have the same effect on demand irrespective of how that growth is distributed across regions. Alternatively, the hierarchical model captures regional income elasticitiesand can therefore be used to analyze the impact of a change in the relative income across regions on material consumption.

2.3 | Shrinkage in hierarchical models

To validate the benefit of using a hierarchical structure in materials demand models, we compare the posterior estimates from hierarchical in comparison to those from a non-hierarchical model using shrinkage plots. When local groups are homogeneous, hierarchical models can account for this

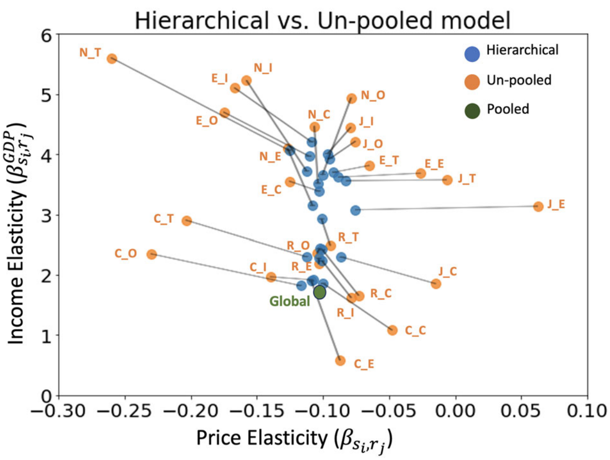


FIGURE 3 Mean values of income elasticity $\beta^{GDP}_{s_i,r_j}$ are plotted on the *y*-axis while mean values of price elasticity (β_{s_i,r_j}) are on the *x*-axis. Each point represents a region–sector combination, with blue values estimated from the hierarchical model and orange values estimated the un-pooled model. A line connects the values estimated from both models for a particular region–sector combination. Following the line from orange point to blue point maps the change in the mean value of $\beta^{GDP}_{s_i,r_j}$ and β_{s_i,r_j} when using a hierarchical model to partially pool data. Region–sector combinations are represented by initials R_S (R is the first letter of the five regions: **C**hina, **E**urope, **N**AM, **J**apan, and **R**oW. S is the first letter of the five sectors: **C**onstruction, **E**lectrical, **T**ransportation, **I**ndustrial, and **O**ther). The average global price and income elasticity estimate from the pooled model is shown by the green point. Data used to make this figure are in Supporting Information Table S1. The shrinkage plot for the intercept and price is in Supporting Information Figure S2.

information by pulling parameters toward a global mean. The shrinkage effect comes from information on the parameter estimate for a particular region–sector combinationA impacting the global mean value, which in turn impacts the value for the parameter estimate in another region–sector B. "Shrinkage" or "pooling" is larger when there is homogeneity in estimates across sites and shrinkage is lesser when sites are heterogeneous.

Figure 3 shows the shrinkage in the mean values for price and income elasticities. There is large shrinkage for the price elasticities, that is, all values are pulled close to the global mean of 0.104 in the hierarchical model. The variation in the price elasticities estimated by the un-pooled model is likely a function of noise rather than structural differences between regions and sectors. However, despite shrinkage caused by pooling, we observe that the hierarchical model finds heterogeneity in income elasticities (observed by the spread along the *y*-axis). The variation across sites in the un-pooled model remains after pooling means, implying that there are likely structural differences across income elasticities of different regions. China and RoW have lower average values for income elasticities than Europe, NAM, and Japan. The fully pooled model only infers a "global" value of income and price elasticity (green point in Figure 2) and is unable to capture these regional differences. Moreover, the estimated "global" income elasticity value is close to the income elasticity values for China, likely because the estimate is dominated by China (which is the largest consumer of copper). The global elasticity from the pooled model is biased by the largest historical consumer and will not give accurate demand predictions for future scenarios in which the relative consumption between regions changes.

The hierarchical model can be considered as a useful hybrid between the un-pooled model which captures regional heterogeneity, and the fully pooled model that has lower uncertainty. In Supporting Information Figure S1, we can also see a strong shrinkage effect for the intercept signifying that there is little difference across sites.

2.4 | Uncertainty reduction for predicted future demand

We sample parameters from their posterior distribution to predict demand in a constant price scenario (prices fixed at 2014 levels; details in Supporting Information Section G). We explore the sensitivity of this prediction under different price scenarios in Supporting Information Section G.

Total copper consumption under a constant price scenario for the hierarchical model was 32.5 and 39 Mt in 2018 and 2040, respectively, where the 2040 demand estimate reflects that of Elshkaki et al. (2016), which estimated 42 Mt copper consumption under market-first and policy-first UNEP GEO-4 foundational scenarios in 2040. While our result is a decent approximation of future copper demand, the focus of this analysis is not on accurately predicting demand but rather using a simple model to describe the benefit of hierarchical pooling especially when there are limited data. Analyses that aim to accurately predict copper demand should use more data and account for shifts in various end-use sectors such as the effect of a transition to renewable energy (see Discussion section).

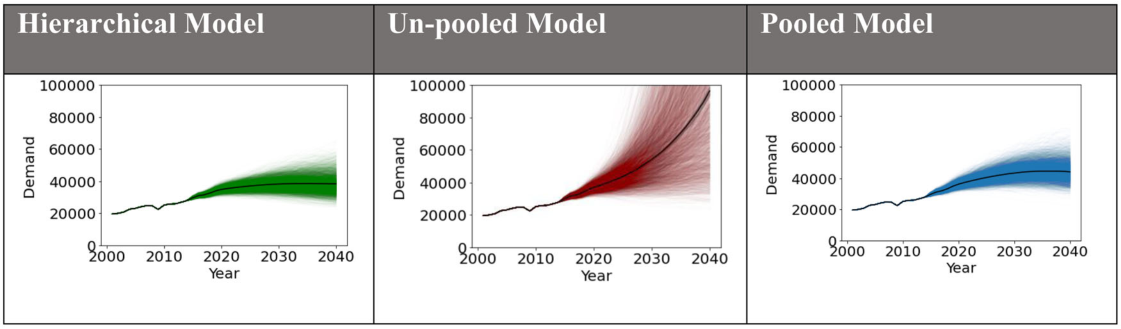


FIGURE 4 Predicted values for future global copper demand based on randomly sampling parameter 2000 values from the estimated posterior probability distributions. Each colored line represents a particular sampling run of the model, with the prediction year on *x*-axis and annual demand on *y*-axis. Three model results shown here: (L) hierarchical model, (C) un-pooled model, (R) fully pooled model. Data used to make prediction are present in Supporting Information Section G.

TABLE 2 Predicted values for demand (in kt) for copper in 2040. Rows represent different indicators for the range of predicted demand values. Columns are different model specifications. Data used to make prediction are present in Supporting Information Section G

Demand @ 2040 (kt)	Hierarchical	Un-pooled	Pooled
Mean	38,570	155,000	44,741
Median	37,933	92,000	44,208
SD	6234	267,244	6712
5th Percentile	29,624	38,662	34,770
25th Percentile	34,287	60,990	39,991
75th Percentile	42,000	156,674	48,708
95th Percentile	49,940	441,100	56,481

The un-pooled model has a poor predictive performance (Figure 4), with the interquartile range for demand in 2040 being 61,000 and 156,000 kt (Table 2). This is driven by the large uncertainty from China intercept ($\beta^{o}_{s_i, r_{China}}$). While the 95% confidence interval includes a value of zero for the China-Electrical intercept ($\beta^{o}_{s_{elec}, r_{China}}$), the intercept has a mean value of 0.10 which means that demand in this sector has about annual growth rate of 10% on average. These uncertainties would lead to a prediction with unrealistic, extreme demand growth.

In comparison, the hierarchical model has much lower uncertainty of prediction (Table 2). The interquartile range for 2040 demand from the hierarchical model is between 34,287 and 42,000 kt (this range is 11.25× lower than the un-pooled model). The shrinkage effect of hierarchical models pools the values of parameters closer by leveraging information across sites and reduces the range of parameter values. While a 95% confidence interval range between 29 and 49 kt is still a large uncertainty for demand prediction, such level of uncertainty may be inevitable given that we specifically chose a case study with small data.

Compared to a fully pooled model which estimates only three parameters (intercept, price elasticity, and income elasticity at a global level), one would expect the hierarchical model to have greater uncertainty as it estimates 81 parameters (local and global effects). However, we find that the hierarchical model has a predictive accuracy similar to, or even better than, the fully pooled model (Table 2). The interquartile range for hierarchical model is 7713 kt (compared to 7717 kt for pooled model) and the SD for hierarchical model is 6234 kt (while it is 6712 kt for fully pooled model).

3 | DISCUSSION

Given the importance of material demand projections to decision-makers, we need novel methods to accurately model material demand at an appropriate level of regional and sectoral detail. We have demonstrated that hierarchical models can (a) reduce uncertainty in demand prediction relative to un-pooled models, while (b) identifying site-specific heterogeneity that fully pooled models cannot. The uncertainty reduction is useful for multiple reasons:

Mine expansion and investment decisions need reliable predictions of long-term demand. Without reliable prediction tools, decision-makers could fail to manage risk and underinvest, leading to supply shortages and price increases for critical minerals. The reduction in uncertainty can also help improve environmental impact estimates, which typically rely on material demand projections as inputs (Van der Voet et al., 2019). Moreover, understanding heterogeneity in price and income elasticities across regions and sectors is important for many decision-makers. There is evidence that developed and developing countries have different metal intensities (Ayres et al., 2003; Krausmann et al., 2009). Having estimates at a local level allows researchers to study various future scenarios for growth in specific regions and sectors with more confidence. One can study, for example, the effect of a GDP slowdown in China on global copper demand without relying on aggregate values for income elasticity.

The fact that the hierarchical model can infer *both* "local" and "global" effects is particularly useful. Pei and Tilton calculate income and price elasticities for 5 developing and 13 developed countries using individual linear regressions (un-pooled) (Pei & Tilton, 1999). To comment on elasticity at an aggregated level, they take the unweighted average of each countries' elasticity value. The hierarchical model improves this in two ways: First, by estimating group-level effects endogenously, we can not only calculate the mean, but also uncertainty in the aggregated global elasticities. Second, the hierarchical model uses the group-level mean to inform the local values, reducing uncertainty. In Pei and Tilton's linear regression analysis for copper demand, only 4 out of the 18 countries had a copper price elasticity value statistically different from 0 (Pei & Tilton, 1999). Similarly, Fernandez (2018) also calculated 174 local price elasticity values and found that only 23% of the estimated by were statistically significant at a p-value of 0.10. In our un-pooled model, which is similar to the approaches in the papers mentioned above, we see a similar result that only 4 out of 25 price elasticity values had a 95% confidence interval that did not include 0. However, in the hierarchical model all of the 25 price elasticities had a 95% confidence interval regime below 0, due to the effects of pooling and shrinkage.

The main drawback of using a Bayesian hierarchical approach compared to traditional regression analyses is that sometimes Bayesian models (especially ones with many parameters) do not converge easily. If sampling chains do not converge, we cannot trust the modeled parameter estimates to be the "true" estimate. In the absence of convergence, the model may not have evaluated all the potential values for that parameter. Increasing the size of the sampling chain and giving the model more time to converge can overcome this challenge but may lead to longer estimation times. If the metal demand module is a part of a larger sustainability or climate model, long estimation times can be undesirable. We did not encounter convergence issues in our analysis (see Supporting Information Section E for details about model run time, hardware, and divergences).

Another limitation of this paper is that we only had access to data from 2000 to 2014 for copper demand that was disaggregated at both a regional and a sectoral level. While using a small dataset was useful in meeting the goals of this paper (i.e., introducing the new methodology and demonstrating how it reduces uncertainty in low-data settings), it limited the analyses we could do. Future studies can extend this model in many ways to conduct interesting research on more recent and detailed datasets. First, while we assumed that local parameters were sampled *independently* from a global distribution of priors, future researchers can study correlations between elasticity estimates in different sites (a full covariance matrix similar to Wang et al. would increase the number of parameters to 625, which is more than the number of datapoints we had). Second, future research should extend this methodology to study the effects of covariates other than price and income. For example, researchers can study the effect of sector-specific variables (e.g., automotive sales, electricity demand), region-specific variables (e.g., urbanization), or recycling rates, which impacts demand through feedback (Ryter et al., 2021, 2022). Adding more covariates can allow interesting analyses such as studying the impact of vehicle electrification (by using EV sales as an inflow variable that impacts demand), transition to renewable energy, and increased recycling rates on metal consumption. Third, future work should apply this model to study demand for critical metals other than copper and shed light on differences across metals markets. Models that quantify the heterogeneity in how metal demand in each industry responds to income and technological changes are key to informing substitution and circular economy strategies (Seck et al., 2020). Demand for many metals can be modeled together by adding another level of hierarchy (e.g., the income elasticity for copper, iron, and aluminum can be sampled from a distribution of the income elasticity for metals). Finally, as also noted by Lupton and Allwood (2018) who apply Bayesian inference to reduce uncertainty in materials flow analyses, a Bayesian approach can improve stock dynamics driven models (e.g., by reducing uncertainty in metal intensity-of-use parameters).

ACKNOWLEDGMENTS

The authors would like to acknowledge Prof. Tamara Broderick and Brian Trippe for introducing them to the world of Bayesian modeling and inference. The authors took help and feedback from many individuals while writing up this paper and would like to acknowledge the help of Nicolas Rothbacher, John Ryter, Xinkai Fu, and Prof. Robert Townsend.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supporting information of this article.

ORCID

Karan Bhuwalka https://orcid.org/0000-0002-1963-6717
Elizabeth A. Moore https://orcid.org/0000-0003-4381-391X
Randolph E. Kirchain https://orcid.org/0000-0002-4019-8598
Elsa A. Olivetti https://orcid.org/0000-0002-8043-2385

REFERENCES

Ayres, R. U., Ayres, L., & Råde, I. (2003). *The life cycle of copper, its co-products and byproducts*. Kluwer Academic.

Bader, H.-P., Scheidegger, R., Wittmer, D., & Lichtensteiger, T. (2011). Copper flows in buildings, infrastructure and mobiles: A dynamic model and its application to Switzerland. *Clean Technologies and Environmental Policy*, *13*(1), 87–101. https://doi.org/10.1007/s10098-010-0278-4

Börjeson, L., Höjer, M., Dreborg, K.-H., Ekvall, T., & Finnveden, G. (2006). Scenario types and techniques: Towards a user's guide. *Futures*, *38*(7), 723–739. https://doi.org/10.1016/j.futures.2005.12.002

Ciacci, L., Fishman, T., Elshkaki, A., Graedel, T. E., Vassura, I., & Passarini, F. (2020). Exploring future copper demand, recycling and associated greenhouse gas emissions in the EU-28. *Global Environmental Change*, *63*, 102093. https://doi.org/10.1016/j.gloenvcha.2020.102093

Crompton, P. (2015). Explaining variation in steel consumption in the OECD. *Resources Policy*, *45*, 239–246. https://doi.org/10.1016/j.resourpol.2015.06.005

Deetman, S., Pauliuk, S., van Vuuren, D. P., van der Voet, E., & Tukker, A. (2018). Scenarios for demand growth of metals in electricity generation technologies, cars, and electronic appliances. *Environmental Science & Technology*, *52*(8), 4950–4959. https://doi.org/10.1021/acs.est.7b05549

Dhar, S., Pathak, M., & Shukla, P. R. (2020). Transformation of India's steel and cement industry in a sustainable 1.5°C world. *Energy Policy*, *137*, 111104. https://doi.org/10.1016/j.enpol.2019.111104

Dong, D., Tukker, A., & Van der Voet, E. (2019). Modeling copper demand in China up to 2050: A business-as-usual scenario based on dynamic stock and flow analysis. *Journal of Industrial Ecology*, *23*(6), 1363–1380. https://doi.org/10.1111/jiec.12926

Dong, J., Liao, J., Huan, X., & Cooper, D. (2022). Expert Elicitation and Data Noise Learning for Material Flow Analysis using Bayesian Inference. arXiv preprint arXiv:2207.09288.

Elshkaki, A., Graedel, T. E., Ciacci, L., & Reck, B. K. (2016). Copper demand, supply, and associated energy use to 2050. *Global Environmental Change*, *39*, 305–315. https://doi.org/10.1016/j.gloenvcha.2016.06.006

Elshkaki, A., Graedel, T. E., Ciacci, L., & Reck, B. K. (2018). Resource demand scenarios for the major metals. *Environmental Science and Technology*, *52*, 2491–2497. https://doi.org/10.1021/acs.est.7b05154

Elshkaki, A., Lei, S., & Chen, W.-Q. (2020). Material-energy-water nexus: Modelling the long term implications of aluminium demand and supply on global climate change up to 2050. *Environmental Research*, *181*, 108964. https://doi.org/10.1016/j.envres.2019.108964

Evans, M., & Lewis, A. C. (2005). Dynamic metals demand model. *Resources Policy*, *30*(1), 55–69. https://doi.org/10.1016/j.resourpol.2004.12.003

Fernandez, V. (2018). Price and income elasticity of demand for mineral commodities. *Resources Policy*, *59*, 160–183. https://doi.org/10.1016/j.resourpol.2018.06.013

Fisher, F. M., Cootner, P. H., & Baily, M. N. (1972). Econometric Model of the World Copper Industry. *Bell Journal of Economics and Management Science*, *3*(2), 568–609.

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *182*(2), 389–402. https://doi.org/10.1111/rssa.12378

Gelman, A. (2006). Multilevel (hierarchical) modeling: What it can and cannot do. *Technometrics*, *48*(3), 432–435. https://doi.org/10.1198/004017005000000661

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/b16018

Gerst, M. D. (2009). Linking material flow analysis and resource policy via future scenarios of in-use stock: An example for copper. *Environmental Science and Technology*, *43*(16), 6320–6325. https://doi.org/10.1021/es900845v

Glöser, S., Soulier, M., & Espinoza, L. A. T. (2013). Dynamic analysis of global copper flows. Global stocks, postconsumer material flows, recycling indicators, and uncertainty evaluation. *Environmental Science and Technology*, *47*(12), 6564–6572. https://doi.org/10.1021/es400069b

Graedel, T. E., Harper, E. M., Nassar, N. T., Nuss, P., & Reck, B. K. (2015). Criticality of metals and metalloids. *Proceedings of the National Academy of Sciences*, *112*(14), 4257–4262. https://doi.org/10.1073/pnas.1500415112

He, R., & Small, M. J. (2022). Forecast of the U.S. copper semand: A framework based on scenario analysis and stock dynamics. *Environmental Science & Technology*, *56*(4), 2709–2717. https://doi.org/10.1021/acs.est.1c05080

Kapur, A., & Graedel, T. E. (2006). Copper mines above and below the ground. *Environmental Science & Technology*, *40*(10), 3135–3141. https://doi.org/10.1021/es0626887

Klinger, J. M. (2020). Environment, development, and security politics in the production of Belt and Road spaces. *Territory, Politics, Governance*, *8*(5), 657–675. https://doi.org/10.1080/21622671.2019.1582358

Krausmann, F., Gingrich, S., Eisenmenger, N., Erb, K.-H., Haberl, H., & Fischer-Kowalski, M. (2009). Growth in global materials use, GDP and population during the 20th century. *Ecological Economics*, *68*(10), 2696–2705. https://doi.org/10.1016/j.ecolecon.2009.05.007

Kuipers, K. J. J., van Oers, L. F. C. M., Verboon, M., & van der Voet, E. (2018). Assessing environmental implications associated with global copper demand and supply scenarios from 2010 to 2050. *Global Environmental Change*, *49*, 106–115. https://doi.org/10.1016/j.gloenvcha.2018.02.008

Leader, A., Gaustad, G., & Babbitt, C. (2019). The effect of critical material prices on the competitiveness of clean energy technologies. *Materials for Renewable and Sustainable Energy*, *8*(2), 8. https://doi.org/10.1007/s40243-019-0146-z

Liu, G., Bangs, C. E., & Müller, D. B. (2013). Stock dynamics and emission pathways of the global aluminium cycle. *Nature Climate Change*, *3*(4), 338–342. https://doi.org/10.1038/nclimate1698

Lupton, R. C., & Allwood, J. M. (2018). Incremental material flow analysis with Bayesian inference. *Journal of Industrial Ecology*, *22*(6), 1352–1364. https://doi.org/10.1111/jiec.12698

Månberger, A., & Stenqvist, B. (2018). Global metal flows in the renewable energy transition: Exploring the effects of substitutes, technological mix and development. *Energy Policy*, *119*, 226–241. https://doi.org/10.1016/j.enpol.2018.04.056

Meager, R. (2019). Understanding the average impact of microcredit expansions: A Bayesian hierarchical analysis of seven randomized experiments. *American Economic Journal: Applied Economics*, *11*(1), 57–91. https://doi.org/10.1257/app.20170299

Müller, E., Hilty, L. M., Widmer, R., Schluep, M., & Faulstich, M. (2014). Modeling metal stocks and flows: A review of dynamic material flow analysis methods. *Environmental Science & Technology*, *48*(4), 2102–2113. https://doi.org/10.1021/es403506a

Nassar, N. T., Barr, R., Browning, M., Diao, Z., Friedlander, E., Harper, E. M., Henly, C., Kavlak, G., Kwatra, S., Jun, C., Warren, S., Yang, M.-Y., & Graedel, T. E. (2012). Criticality of the geological copper family. *Environmental Science & Technology*, *46*(2), 1071–1078. https://doi.org/10.1021/es203535w

Pei, F., & Tilton, J. E. (1999). Consumer preferences, technological change, and the short-run income elasticity of metal demand. *Resources Policy*, *25*(2), 87–109. https://doi.org/10.1016/S0301-4207(99)00013-6

Ryter, J., Fu, X., Bhuwalka, K., Roth, R., & Olivetti, E. A. (2021). Emission impacts of China's solid waste import ban and COVID-19 in the copper supply chain. *Nature Communications*, *12*(1), 3753. https://doi.org/10.1038/s41467-021-23874-7

Ryter, J., Fu, X., Bhuwalka, K., Roth, R., & Olivetti, E. (2022). Assessing recycling, displacement, and environmental impacts using an economics-informed material system model. *Journal of Industrial Ecology*, *26*, 1010–1024. https://doi.org/10.1111/jiec.13239

Schipper, B. W., Lin, H.-C., Meloni, M. A., Wansleeben, K., Heijungs, R., & van der Voet, E. (2018). Estimating global copper demand until 2100 with regression and stock dynamics. *Resources, Conservation and Recycling*, *132*, 28–36. https://doi.org/10.1016/j.resconrec.2018.01.004

Seck, G. S., Hache, E., Bonnet, C., Simoën, M., & Carcanague, S. (2020). Copper at the crossroads: Assessment of the interactions between low-carbon energy transition and supply limitations. *Resources, Conservation and Recycling*, *163*, 105072. https://doi.org/10.1016/j.resconrec.2020.105072

S&P Global Market Intelligence. (2016). *S&P Capital IQ*. https://www.snl.com

Stuermer, M. (2017). Industrialization and the demand for mineral commodities. *Journal of International Money and Finance*, *76*, 16–27. https://doi.org/10.1016/j.jimonfin.2017.04.006

United Nations. (2010). World bank open data initiative. https://data.worldbank.org

Van der Voet, E., Van Oers, L., Verboon, M., & Kuipers, K. (2019). Environmental implications of future demand scenarios for metals: Methodology and application to the case of seven major metals. *Journal of Industrial Ecology*, *23*(1), 141–155. https://doi.org/10.1111/jiec.12722

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Vidal, O., Le Boulzec, H., & François, C. (2018). Modelling the material and energy costs of the transition to low-carbon energy. *EPJ Web of Conferences*, *189*, 00018. https://doi.org/10.1051/epjconf/201818900018

Wang, T., Berrill, P., Zimmerman, J. B., Rao, N. D., Min, J., & Hertwich, E. G. (2022). Improved copper circularity as a result of increased material efficiency in the U.S. housing stock. *Environmental Science & Technology*, *56*, 4656–4577. https://doi.org/10.1021/acs.est.1c06474

Wang, S., Sun, X., & Lall, U. (2017). A hierarchical Bayesian regression model for predicting summer residential electricity demand across the U.S.A. *Energy*, *140*, 601–611. https://doi.org/10.1016/j.energy.2017.08.076

Watari, T., Nansai, K., Giurco, D., Nakajima, K., McLellan, B., & Helbig, C. (2020). Global metal use targets in line with climate goals. *Environmental Science & Technology*, *54*(19), 12476–12483. https://doi.org/10.1021/acs.est.0c02471

Watari, T., Nansai, K., & Nakajima, K. (2021). Major metals demand, supply, and environmental impacts to 2100: A critical review. *Resources, Conservation and Recycling*, *164*, 105107. https://doi.org/10.1016/j.resconrec.2020.105107

Zhang, X., Dong, Q., Costa, V., & Wang, X. (2019). A hierarchical Bayesian model for decomposing the impacts of human activities and climate change on water resources in China. *Science of The Total Environment*, *665*, 836–847. https://doi.org/10.1016/j.scitotenv.2019.02.189

Zhang, L., Yang, J., Cai, Z., & Yuan, Z. (2015). Understanding the spatial and temporal patterns of copper in-use stocks in China. *Environmental Science & Technology*, *49*(11), 6430–6437. https://doi.org/10.1021/acs.est.5b00917

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Bhuwalka, K., Choi, E., Moore, E. A., Roth, R., Kirchain, R. E., & Olivetti, E. A. (2022). A hierarchical Bayesian regression model that reduces uncertainty in material demand predictions. *Journal of Industrial Ecology*, 1–13. https://doi.org/10.1111/jiec.13339