Mapping the material stock spatial distribution from thin-film CdTe utility-scale photovoltaic arrays in the United States

Mia Dagati, Jacob T. Stid, Annick Anctil,

Michigan State University

East Lansing, Michigan 48823, United States

***Abstract* — Understanding the material composition of utility-scale cadmium telluride (CdTe) photovoltaic fleets is fundamental to effective end-of-life planning and recycling infrastructure investment. This study develops a classification framework to identify First Solar module generations and quantify embodied materials at the array level across the contiguous United States, drawing on spatial and geometric metadata from the GM-SEUS v2.0 dataset. A year-anchored decision rule assigned 456 arrays to product generations spanning Series 2 through 7, with per-array module counts derived from a tiered geometry estimation approach and material quantities calculated using generation-specific intensity factors. The resulting inventory captures roughly 133 million modules representing 19.2 GW of installed capacity, with embedded stocks of 2,458 metric tons of CdTe, 136,725 metric tons of aluminum, 19,255 metric tons of steel, 28,363 metric tons of EVA encapsulant, and 395 metric tons of copper. Aggregate capacity proves a poor indicator of regional material exposure, as module counts, array tallies, and material shares diverge considerably across generations and geographies. These findings offer a higher-resolution basis for anticipating secondary material flows and evaluating where recycling capacity may fall short of future demand.**

I. INTRODUCTION

The United States has experienced unprecedented growth in solar energy deployment over the past two decades, driven by declining technology costs, federal investment incentives, and state-level renewable portfolio standards~~.~~[1]. As of the mid-2020s, utility-scale solar accounts for a substantial and growing share of national electricity generation capacity, with installations spanning a wide range of climatic regions, land use contexts, and technology types[2]. Understanding the material composition and lifecycle implications of these systems has become increasingly important as the first generations of utility-scale arrays approach end-of-life and policymakers, developers, and recyclers begin assessing the scale of future waste pipelines.

Among utility-scale PV technologies, thin-film cadmium telluride (CdTe) has emerged as a dominant technology for large-scale ground-mounted deployment in the United States, mainly due to the commercial success of First Solar, the primary manufacturer of CdTe modules~~.~~[3]. CdTe systems offer several advantages at utility scale, including lower manufacturing energy requirements, strong performance under high-temperature, higher energy efficiencies, and slower degradation rates than current silicone technologies~~.~~[4]. These characteristics have made CdTe a preferred technology for large desert and agricultural installations in states such as California, Texas, and Florida, where the combination of available land, solar resource, and transmission infrastructure has enabled deployment at gigawatt scale.

Despite this prominence, CdTe technology presents unique material management challenges. Cadmium is a regulated toxic heavy metal, and tellurium is a rare and economically critical element with limited global supply. With both materials concentrated in a small number of large plants, end-of-life decisions at individual sites carry significant consequences for material recovery, environmental liability, and secondary supply. At the same time, the aluminum and steel used in module frames and mounting systems represent a substantial quantity of recyclable resources that could properly recovered and routed to appropriate facilities.

Planning for these end-of-life flows requires knowing not just how many modules are currently installed, but what specific module series and where they are installed , since material composition varies substantially across First Solar's product generations. First Solar has commercialized multiple CdTe module series over its history, from the early Series 2 through the current Series 7, with each generation reflecting advances in module efficiency and material use. The transition from earlier to later series represents a significant shift, with newer modules being substantially larger in area and incorporating different structural materials, meaning that the series composition of a regional array directly determines the type and quantity of materials that will reach end-of-life simultaneously.

Large geospatial datasets of solar infrastructure provide detailed information on the location, capacity, and footprint of utility-scale arrays. However, only the Ground-Mounted Solar Energy in the United States (GM-SEUS) dataset provides continuous sub-array-level information on installation layout across the United States[5]. According to GM-SEUS v2.0, 456 CdTe arrays are identified nationwide, with these systems typically deployed at very large scales, concentrating a substantial portion of total utility-scale capacity within a relatively small number of installations. However, GM-SEUS does not identify module manufacturer or series, limiting its direct use for estimating material composition and creating a gap between what is spatially documented and what is needed for material stock and recycling planning.

This study addresses that limitation by developing a data-driven framework to infer CdTe module series and estimate embodied material inventories at the array level across the contiguous United States, providing a foundation for evaluating current material stocks and informing future end-of-life management strategies. The overall study workflow, spanning data integration, hybrid series classification, validation, material composition factor extraction, module inventory calculation, spatial aggregation and mapping, trend analysis, and recycling infrastructure recommendations, is summarized in Figure 1.

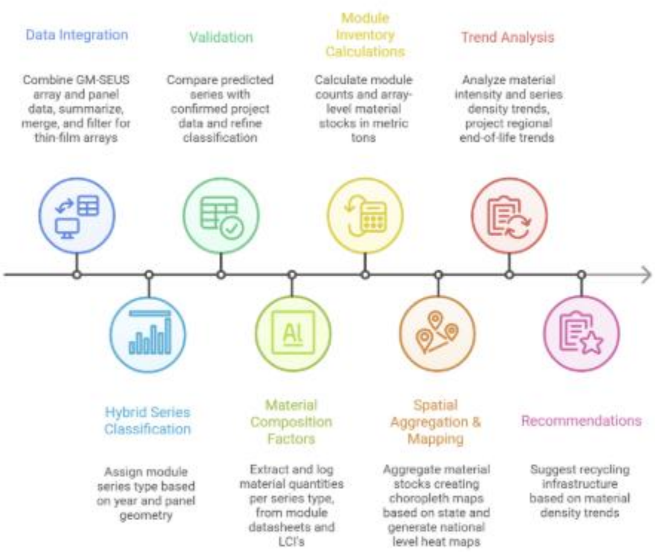

Fig. 1 Quantifying and Analyzing Material Intensities of Thin-Film Utility Level Solar Arrays Workflow

II. METHODS

We developed a data-integration framework to classify First Solar cadmium telluride (CdTe) module series (Series 2–7) and estimate embodied material inventories for utility-scale photovoltaic arrays across the contiguous United States. Array- and panel-level metadata from GM-SEUS, including total row area, average row width, row length, mounting configuration, and orientation, were imported as separate tables, standardized, and merged into a unified array-based dataset. In the panel-row dataset, placeholder values of −9999 were treated as missing, and row-level geometry outliers were removed using a three standard deviation threshold before aggregating panel metrics by array. The dataset was then filtered to thin-film arrays within the contiguous United States, with arrays linked to state boundaries using Federal Information Processing Standard identifiers.

Module series were assigned using a year-first, rule-based classifier in which arrays were mapped to Series 2, 3, 4, 6, or 7 based on discrete installation-year ranges corresponding to known deployment periods. Years falling within transition periods between series were treated as boundary cases and assigned default labels based on observed deployment timing. Geometric variables were evaluated at boundary periods but were not found to reliably distinguish between adjacent series and therefore did not override year-based assignments. A year confidence metric was calculated by comparing reported installation year against an Landsat-derived estimate and categorized as high, medium, or low based on the magnitude of disagreement, with 52.2% of arrays receiving high confidence scores and 23.7% rated low confidence. The series-specific deployment year ranges and boundary period assignments are summarized above in Figure 2.

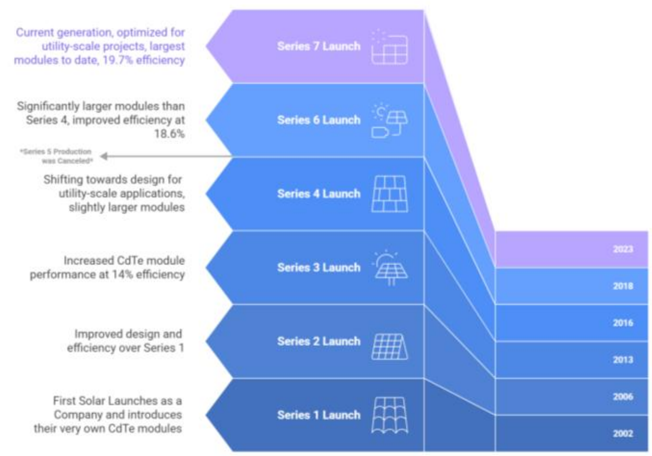

Fig. 2 First Solar Series Production Timeline

Module counts were estimated using a hierarchical approach to module area determination. Panel-derived average row width and length were used when available, array-level geometry fields served as a secondary source, and series-specific module area values from manufacturer datasheets were applied as a fallback. Of the 456 classified arrays, 55.9% relied on this specification sheet fallback rather than observed panel- or array-level geometry, reflecting incomplete coverage of panel-level geometry fields in GM-SEUS. While manufacturer-reported nominal dimensions are considered reliable central estimates, this fallback rate introduces some uncertainty for a portion of arrays, as it does not capture variation in installation practices, module orientation, or non-standard row configurations, providing instead a general estimate of module count per array. At the scale of the overall dataset, however, this margin of uncertainty is small and has a negligible impact on aggregate results. The three-tier hierarchy and the share of arrays falling into each tier is illustrated in Figure 3

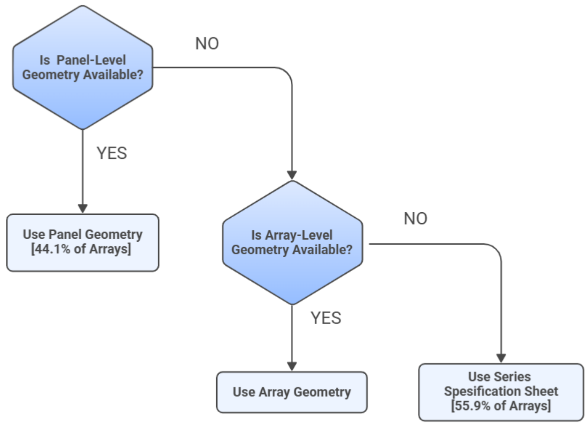

Fig. 3 Decision Nodes for Methods of Estimating Module Quantity per Array

Module counts were then multiplied by fixed, series-specific material intensity factors, including CdTe, steel, aluminum, ethylene-vinyl acetate (EVA), copper, and other components, derived from manufacturer datasheets and peer-reviewed literature to estimate array-level embodied material inventories. All processing was conducted in R using tidyverse-based workflows, with geospatial analysis performed using the sf and tigris packages.

III. RESULTS

Across the contiguous United States, the framework characterized 456 thin-film CdTe arrays comprising approximately 133 million modules with a combined installed capacity of 19.2 GW. Total estimated embodied materials include 2,458 metric tons of CdTe, 136,725 metric tons of aluminum, 19,255 metric tons of steel, 28,363 metric tons of EVA encapsulant, and 395 metric tons of copper.

Although Series 4 accounted for the largest share of classified arrays (36.2%), it did not lead in total module count, installed capacity, average per array capacity or module efficiency across all series. Series 3 represented the greatest number of individual modules nationally, at approximately 39.0 million modules, reflecting its deployment in large-capacity installations with high module counts despite fewer discrete arrays. Series 3 arrays averaged 338,992 modules per array and 40.2 MW per array, the largest average module count of any series, indicating that Series 3 was predominantly deployed in very large utility-scale configurations. Series 2 arrays averaged 195,411 modules and 28.7 MW per array, consistent with the earlier period of utility-scale CdTe deployment when per-project capacities had not yet reached the scale seen in later series. Series 4 arrays averaged 297,263 modules and 26.9 MW per array, suggesting that instillations were on average deployed at lower per-project capacities than series 3. The distribution of arrays, modules, and capacity by series is shown in Figure 4.

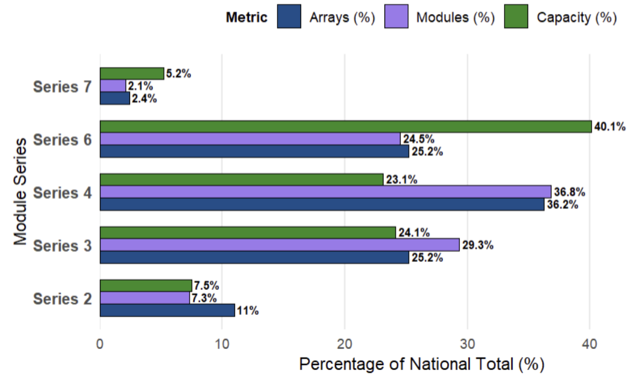

Fig. 4 Distribution of Arrays, Modules, and Installed Capacity by Module Series

The transition to the larger-format Series 6 and 7 modules is reflected clearly in the array-scale metrics. Series 6 arrays averaged 283,075 modules and 66.9 MW per array, while Series 7 arrays averaged 251,145 modules and 91.6 MW per array, the highest average capacity of any series. The substantially larger module area of Series 6 and 7 compared to Series 2 through 4 means that fewer modules are needed to achieve the same row area coverage, and the resulting arrays are substantially larger in capacity, driven primarily by per-module area increases rather than efficiency gains. This generational shift in module format has direct implications for material intensity: Series 6 dominates national aluminum inventory entirely, contributing 136,725 metric tons, while Series 7 accounts for all estimated national steel inventory at 19,255 metric tons, with Series 2 through 4 contributing

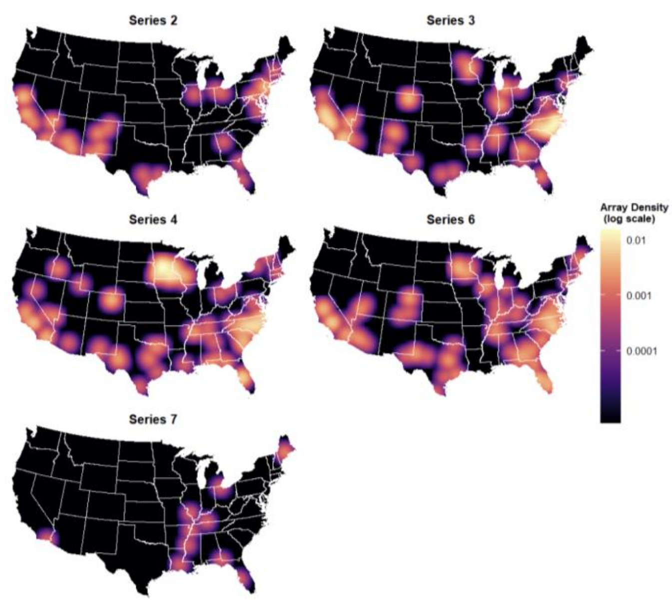

Fig. 5 National Density map of arrays by First Solar series number

neither aluminum nor steel under the current material intensity framework.

Series 6 dominated installed capacity at 40.1% of the national total, a consequence of its larger module format and the scale of recent deployments. This divergence across metrics, where array count, module count, and capacity each point to a different dominant series, illustrates that no single metric adequately characterizes the composition of the national CdTe fleet and reinforces the importance of series-level characterization when estimating material stocks.

Substantial spatial heterogeneity in material distribution was found across the contiguous US, driven largely by regional differences in the age and type of installed module series, as shown in the national density maps in Figure 5. Figure 5 maps array density by product series across the contiguous United States, where each panel represents a single series and the color scale reflects the density of arrays within a given area on a log scale, allowing both high-concentration hotspots and more diffuse deployment patterns to be visible simultaneously. To further illustrate the geographic concentration of individual materials, Figure 6 maps the spatial density of each material type across the contiguous United States. Each panel displays the proportion of that material's total national inventory within a given area, normalized independently so that the maps reflect relative geographic concentration rather than absolute mass. A log scale is applied to each panel to improve visibility of mid-tier density regions that would otherwise be obscured by the dominance of a small number of high-concentration locations.

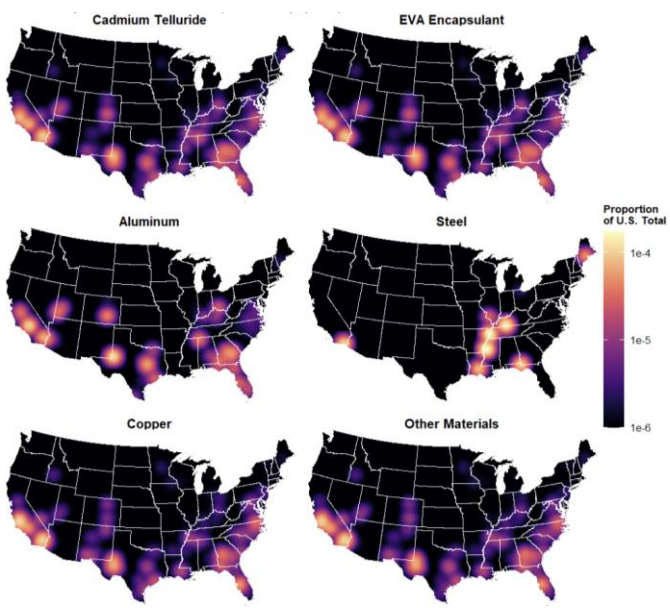

Fig. 6 Proportional Geographic Concentration of Thin-Film Solar Array Materials by Type

California alone accounts for approximately 25.5% of total installed CdTe modules nationwide, with deployment concentrated in the Central Valley and southeastern desert regions, particularly in Kern, Riverside, and Imperial counties. Despite this dominant share of installations, California contributes only 12.3% of national aluminum inventory, reflecting its higher prevalence of Series 3 and 4 modules, with a CdTe intensity of 0.138 t/MW and aluminum intensity of 15.9 t/MW consistent with a mixed-generation fleet. In contrast, Texas represents 18.1% of installed modules but contributes 20.0% of national aluminum inventory, demonstrating that similar module counts can mask substantial differences in embodied material stocks depending on series composition.

Material intensity per unit capacity varied substantially across states beyond the California and Texas comparison. Florida exhibited a steel intensity of 1.199 t/MW driven by its concentration of Series 7 arrays, while Louisiana showed the highest steel intensity of any state at 4.198 t/MW despite a total installed base of 435 MW. Alabama exhibited an aluminum intensity of 19.7 t/MW, the highest among top states, suggesting a high prevalence of Series 6 installations relative to its total capacity of 361 MW. In contrast, Texas, Nevada, Georgia, and Arizona showed zero or near-zero steel intensity, indicating negligible Series 7 deployment. Illinois, with a steel intensity of 1.367 t/MW alongside an aluminum intensity of only 0.173 t/MW, suggests a fleet composition weighted toward Series 7 rather than Series 6, despite a total installed capacity of 444 MW. These intensity differences have direct implications for end-of-life planning, as two states with identical installed capacity can require fundamentally different recycling infrastructure depending on whether their arrays are aluminum-framed Series 6 or steel-framed Series 7 systems. When these material inventories are considered alongside the geographic distribution of existing CdTe recycling facilities, several high-inventory states show limited proximate recycling capacity, pointing to potential infrastructure gaps where end-of-life material availability may outpace local recovery capability.

IV. DISCUSSION AND FUTURE WORK

These results show that relying solely on installed capacity or module counts can substantially misrepresent regional material inventories and future end-of-life flows. The divergence between array count, module count, and capacity shares across series, and between module share and material share across states, demonstrates that series-level characterization is essential for accurate material stock accounting. The generational shift from compact Series 2 through 4 modules to the larger-format Series 6 and 7 has not only changed the scale at which individual arrays are deployed, but has fundamentally altered the type and quantity of structural materials embedded in the national fleet, with aluminum and steel now representing the dominant end-of-life material streams for newer installations.

Because recycling capacity, policy requirements, and material recovery practices vary by region, inaccurate estimates of when and how much material will reach end-of-life availability could lead to poorly sited or improperly sized recycling infrastructure and missed opportunities to recover valuable materials, increasing global warming potential. By characterizing CdTe material stocks at the array and module-series level and comparing projected material flows against existing recycling infrastructure, this study provides a more accurate basis for projecting secondary material supply, identifying infrastructure gaps, and supporting policies that strengthen supply-chain resilience and advance circular-economy goals for utility-scale solar.

Future work could extend this framework to other PV technologies such as crystalline silicon, which represents the majority of utility-scale capacity but has received comparatively less attention in material stock analyses. Incorporating temporal projections of module retirement, based on installation year and assumed service life distributions, would enable estimation of annual material availability rather than cumulative stock, providing a more actionable basis for recycling infrastructure planning. Such projections would illustrate how the timing and geography of material availability will shift as older Series 2 through 4 installations reach end-of-life in the near term while larger Series 6 and 7 arrays follow in subsequent decades, with the regional concentration of those flows having direct implications for where and when recycling capacity will be needed most. Integrating data on existing and planned recycling facilities with these temporal projections would allow direct assessment of regional capacity gaps and inform siting decisions for new infrastructure. Improved remote sensing methods and more complete manufacturer reporting could further reduce uncertainty in module series classification and material intensity estimates, particularly by reducing reliance on specification sheet fallbacks for module area determination.

References

[1] D. Feldman and J. Zuboy, "Quarterly Solar Industry Update: Fall 2024," U.S. Department of Energy, Office of Clean Energy Demonstrations, Washington, DC, USA, Oct. 2024. [Online]. Available:https://www.energy.gov/cmei/systems/quarterly-solar-industry-update

[2] J. Seel, J. M. Kemp, A. Cheyette, D. Millstein, W. Gorman, S. Jeong, D. Robson, R. Setiawan, and M. Bolinger, "Utility-Scale Solar, 2024 Edition: Empirical Trends in Deployment, Technology, Cost, Performance, PPA Pricing, and Value in the United States," Lawrence Berkeley National Laboratory, Berkeley, CA, USA, 2024. [Online]. Available:https://emp.lbl.gov/publications/utility-scale-solar-2024-edition

[3] B. E. Sartor et al., "Roadmap to 100 GWDC DC: Scientific and Supply Chain Challenges for CdTe Photovoltaics," *Joule*, vol. 10, article no. 102235, Elsevier, 2026. DOI: 10.1016/j.joule.2025.102235

[4] M. A. Scarpulla et al., "CdTe-Based Thin Film Photovoltaics: Recent Advances, Current Challenges and Future Prospects," *Solar Energy Materials and Solar Cells*, vol. 255, p. 112289, 2023. DOI: 10.1016/j.solmat.2023.112289

[5] J. T. Stid, A. D. Kendall, A. Anctil, J. Rapp, J. C. Bingaman, D. W. Hyndman, “A harmonized dataset of ground-mounted solar energy in the US with enhanced metadata”, Sci Data, vol. 12, p. 1586, 2025, doi:10.1038/s41597-025-05862-4.