UCERS

User-Centered Solutions for Digital and Sustainable Energy Communities

UCERS Final Public Report

Project Duration: 01.03.2021 – 28.02.2025

Classification: Public

Version: 1.0 (Final)

Publication Date: 27.05.2025

Prepared for: Austrian Research Promotion Agency (FFG)

Distribution: This document is approved for public release.

Authors:

Matthew Clarke[1], Natalie Taupe[1], Edit Parada[1], Frederike Ettwein[1], Christina Huber[2], Markus Fischer[3], Yi Guo[1], Matthias Malottke[1], Michael Oberhofer[4], Martin Schloffer[4], Peter Illich[2], Rupert Wychera[3], Jacob Wöginger[2], William Markl[5], Markus Plank[6], Tobias Hofer[7], Josef Bauchinger[7], Eszter Gunnarsson[1]

[1] FH Technikum Wien
[2] neoom AG
[3] Energy-Climate GmbH
[4] 4ward Energy Research GmbH
[5] Reisenbauer Solutions GmbH
[6] Spitzer GmbH
[7] Fronius International GmbH

Project "UCERS – User-Centered Solutions for Digital and Sustainable Energy Communities" (FFG number 892267) was funded through the Vorzeigeregion Energie 2019 program.

Document History

Version:	Date:	Released by:	Changes:
1.0	27.05.2025	Matthew Clarke	Initial release, 1st publication

Document Signature

This document has been digitally signed by its issuer, Matthew Clarke, on behalf of the UCERS Project Team, using a qualified electronic signature in accordance with EU eIDAS regulations. Any version lacking this digital signature may have been altered and cannot be verified as authentic.

To verify the authenticity and integrity of this document, please open it in Adobe Acrobat Reader or another compatible PDF viewer that supports digital signature validation.

SIGNATURINFORMATIONEN	**Unterzeichner**	Matthew John Clarke
	Datum/Zeit-UTC	2025-05-27T20:12:20+02:00
	Prüfinformation	Informationen zur Prüfung der elektronischen Signatur finden Sie unter: https://www.signaturpruefung.gv.at
Hinweis	Dieses mit einer qualifizierten elektronischen Signatur versehene Dokument hat gemäß Art. 25 Abs. 2 der Verordnung (EU) Nr. 910/2014 vom 23. Juli 2014 ("eIDAS-VO") die gleiche Rechtswirkung wie ein handschriftlich unterschriebenes Dokument.	

Table of Contents

1 Executive Summary 9
2 Introduction 11
3 Engagement and Participatory Approaches 13
4 Socio-Environmental Evaluation of Energy Communities 18
4.1 Sustainable Development 18
4.2 An Integrated Approach to Sustainability 19
4.3 Role of Energy Communities in Sustainable Development 19
4.4 Methodology 20
4.4.1 Development of Evaluation Criteria 22
4.4.2 Feedback on Criteria: Expert Interviews 23
4.4.3 Development and Pre-testing Indicators and Survey 23
4.4.4 Summary of the Evaluation Scheme 24
4.4.5 Application of Evaluation Scheme Nationwide Survey 27
4.4.6 Visualization and Feedback 27
4.4.7 Stakeholder Workshop (Survey Follow-up) 29
4.4.8 Post-Workshop Analysis and Derivation of Recommendations 30
4.5 Results and Discussions 30
4.5.1 Assessment of Energy Communities Using Evaluation Scheme 30
4.5.2 Diverging Practices Across Community Types 33
4.5.3 Self-Assessment and Reflections 33
4.5.4 Identified Goal Conflicts 35
4.6 Recommendations 37
4.7 Conclusions 41
5 Digital Tools for Energy Communities 42
5.1 CDEP 42
5.2 COOP 44
5.2.1 Local Energy Optimization & Community Management 45
5.2.2 Monitoring, Control, and User Interaction 46
5.2.3 Integration of Energy Storage 48
6 Testbed Implementations and Empirical Insights 52
6.1 Selection Criteria for Testbeds 52
6.1.1 neoom: Public and Commercial Buildings with PV and Battery 53
6.1.2 RSO/ENC: Single-Family Homes, PV, Battery, EVs, & Heat Pumps 54
6.1.3 Spitzer: Commercial Buildings with EV Integration 54
6.1.4 Fronius: Apartment Building with PV System 55
6.2 Monitoring and Analysis Methods 55
6.2.1 Metrics and Obtaining Measurements 55
6.2.2 Data Preparation and Validation Approach 58
6.2.3 Community Scenario Analysis 58

6.2.4 Economic Assumptions and Cost Evaluation ... 60
6.3 Results and Discussion ... 60
6.3.1 Limitations of Findings ... 60
6.3.2 Results of Tech-Communities ... 61
6.3.3 Results of Baseline-Communities ... 65
6.4 Conclusions ... 67
7 Project Key Findings ... 68
7.1 Early-Stage Participation and Emerging Expectations ... 68
7.2 Strategic Development and Mature Practices ... 68
7.3 Comparative Insights and Sector Evolution ... 69
7.4 Digital Infrastructure: User Engagement & Optimization ... 69
7.5 Testbeds: Real-World Performance and Lessons ... 70
7.6 Conclusion ... 70
8 Recommendations and Future Perspectives ... 71
8.1 Toward Inclusive and Sustainable ECs ... 71
8.2 Managing Expectations in Early ECs ... 71
8.3 Expanding Access and Inclusion ... 71
8.4 Aligning ECs with Systemic Sustainability Goals ... 72
8.5 Scaling Digital Tools for ECs ... 72
8.6 Smart Design and Inclusive Participation ... 72
8.7 Final Reflections ... 73
8.8 Further Information ... 73
9 Appendices ... 74
9.1 Appendix A: Overview of Survey and Interview Methods ... 74
9.1.1 Quantitative Survey of Members and Prospective Members ... 74
9.1.2 In-Depth Interviews with EC Members ... 76
9.1.3 Stakeholder Interviews: EC Establishment and Operation ... 76
9.1.4 Stakeholder Interviews: Digitalization of ECs ... 77
9.2 Appendix B: Evaluation Scheme ... 78
9.2.1 Description of the Categories and Criteria ... 78
9.2.2 Survey Results ... 90
9.2.3 Survey Tool for Energy Communities ... 99
9.3 Appendix C: Supplemental Information Simulations ... 109
9.3.1 Simulation Methods: Battery Storage ... 109
9.3.2 Results: Simulated Storage ... 111
9.4 Appendix D: Data Validation and Preprocessing ... 114
10 Sources ... 116

From a technical standpoint, the system proved scalable and adaptable. The ability to work with a variety of hardware configurations—without requiring major infrastructure changes—positioned the solution as a flexible option for similar communities. As with other components of the COOP, lessons learned through this deployment informed recommendations for future improvements, including enhanced predictive functions and broader integration with national energy data systems.

The testbed implementations highlighted the ability of the system to deliver individual and aggregated energy data, helping to identify inefficiencies, demand peaks, and optimization potential. This included tools to encourage flexible load shifting—such as timing consumption with periods of low-cost or high-renewable availability—and dashboards to communicate community performance metrics. For further details on technical implementation, system performance, and user feedback, refer to the relevant monitoring sections later in this report, particularly Section 6.

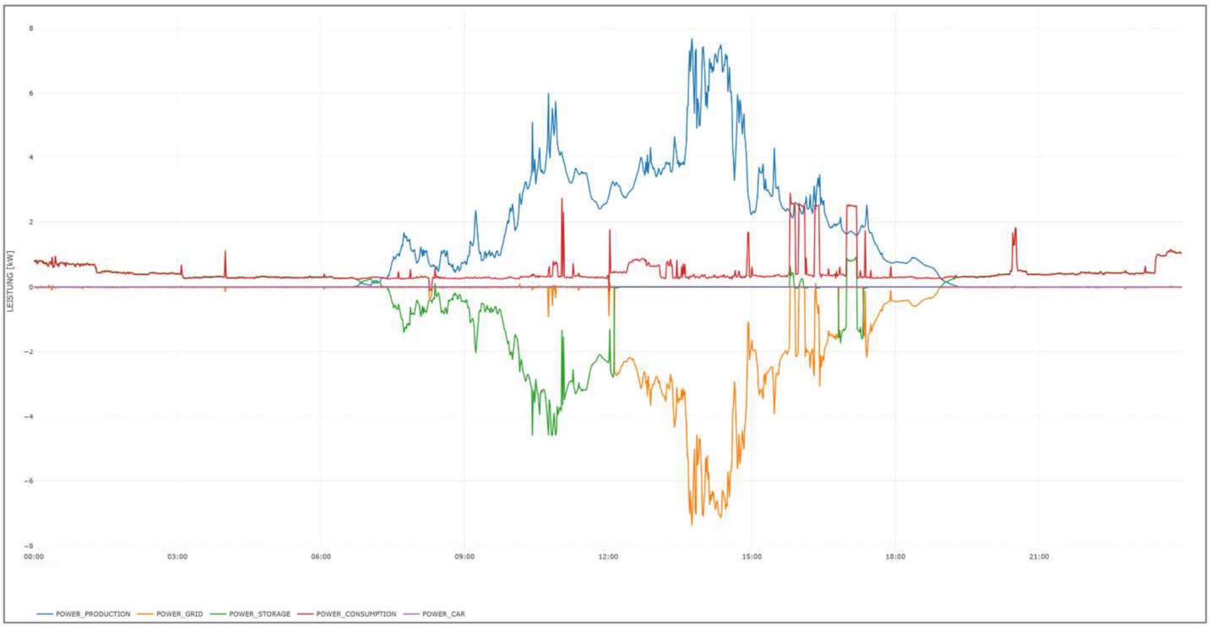

Figure 16: Historical Dashboard Data - Historical data view from the dashboard, showing past energy production, consumption, and storage performance for analysis and optimization purposes (RSO).

5.2.3 Integration of Energy Storage

As part of a dedicated line of work on cooperative energy optimization, the UCERS project investigated how the performance of energy communities could be enhanced through the strategic use of residential storage systems and forecast-based energy planning. These investigations were carried out through a collaboration between 4ward Energy Research and neoom, who combined domain expertise in energy storage systems and data-driven modelling. Their efforts focused on the question of how intelligent storage operation—supported by predictive algorithms—could contribute to increased self-sufficiency, cost efficiency, and coordinated energy use within a community setting.

The analysis built on real consumption data from existing energy communities and tested a range of modelling and optimization strategies, with the aim of informing implementations of advanced energy management practices.

5.2.3.1 Machine Learning for Consumption Forecasting

One part of the investigation focused on the development of machine learning (ML) models to forecast electricity consumption in energy communities. The motivation was to assess whether a “soft sensing” approach—using aggregated smart meter data combined with weather inputs—could provide sufficient accuracy to support storage planning, potentially avoiding the need for widespread installation of real-time measurement infrastructure.

Smart meter data from an energy community comprising 36 mixed residential participants served as the training dataset. The models evaluated included Random Forest Regressors, Neural Networks, and gradient boosting methods such as XGBoost and LightGBM. Weather data (temperature, solar irradiation) and time-related variables were integrated as exogenous inputs.

Initial attempts to predict full 15-minute load profiles proved unreliable due to high variability and the relatively small participant pool. As a result, the team shifted to a simplified approach, forecasting average daytime and nighttime consumption values. This reduced temporal resolution improved model stability and produced more consistent results, particularly for the nighttime period—a time when solar generation ceases, and stored energy becomes a key supply source.

Model accuracy was evaluated using standard error metrics, with results indicating that nighttime forecasts were significantly more reliable, due in part to more predictable user behavior during these hours. The findings suggest that intelligent storage discharge during night periods, guided by such forecasts, could be an effective strategy for maximizing alignment between supply and demand.

5.2.3.2 Simulated Storage Operation for Energy Community Optimization

As part of the UCERS project, a detailed simulation study was conducted to assess how residential battery storage systems can improve energy community performance. The analysis focused on five prosumer households, each equipped with a photovoltaic (PV) system but not yet using a storage unit. For each household, the installation of a 20-kWh battery was simulated under different operational strategies to evaluate the impact on local energy use and economic outcomes. The study examined electricity produced, consumed, and sold, applying a representative tariff structure that included: 29.60 ct/kWh for grid electricity, 18.2 ct/kWh for intra-community consumption, 10 ct/kWh for energy fed into the community, and 6 ct/kWh for feed-in to the public grid (EVU). This pricing framework enabled the evaluation of both individual cost savings and community-level energy performance under various storage operation scenarios.

Three battery storage scenarios were analyzed:

- **SN0: Baseline (no storage)**: Reflects the actual historical state of energy distribution without any storage system in place. This scenario serves as a reference point for evaluating the impact of storage-based optimization strategies.
- **SN1: Self-sufficiency optimization**: Local self-sufficiency optimization (storage used only for household needs). Storage is charged when surplus PV generation is available and discharged when household consumption exceeds generation. Decisions are based purely on energy balance at the individual household level, with the aim of maximizing local self-consumption and minimizing reliance on grid supply.
- **SN2: Self-sufficiency + community optimization** (storage used both locally and to discharge into the community): Builds on the self-sufficiency approach but incorporates economic optimization by considering the higher compensation rate available for energy fed into the energy community. Battery discharge decisions are informed by day-ahead forecasts and are not limited to meeting household demand—energy may also be discharged strategically to supply the community when demand exists, thereby increasing individual revenue.

The results showed that operating storage purely for self-sufficiency reduced the amount of energy available to the broader community. In contrast, adopting a community-oriented optimization approach significantly increased intra-community energy infeed—by as much as a factor of ten—resulting in higher revenues and greater local utilization of renewable energy.

Cost savings for individual prosumers were also higher (28–139 EUR per year) under the full optimization scenario (Self-sufficiency + community optimization), primarily due to the improved compensation rates for energy shared within the community. For example, a representative household (Prosumer C) increased its community infeed by approximately 1,500 kWh per year under community-optimized operation, clearly illustrating the benefits of coordinated storage management.

It should be noted, however, that these results are based on simulations assuming perfect forecasting. A more conservative estimate of the net gain for participating households—accounting for forecasting uncertainty and operational variability—would place the expected benefit in the range of €10 to €60 per year. Nevertheless, with the growing adoption of flexible tariffs by energy providers, these savings could increase in future applications.

At the community level, energy exchange increased by 700 to 3,000 kWh per year, depending on the structure of the system. This improvement in local energy balancing also reduced reliance on the external grid and contributed to more grid-friendly behavior—one of the central aims of modern energy communities.

5.2.3.3 Conclusions and Outlook

The simulation and forecasting activities underline both the potential and the limitations of predictive energy management in community settings. While forecasting household consumption remains challenging due to behavioral variability, solar generation prediction is much more robust—opening the door for forecast-informed storage control as a practical strategy.

Economically, the simulations showed that operating residential storage systems under a traditional self-sufficiency strategy can result in substantial tariff-based savings—approximately €360 to €610 per year, depending on the household's consumption and PV profile. These figures do not account for the cost of the storage system itself, a portion of which would need to be covered by the savings in a full economic assessment. Building on this, the analysis found that adding a community-optimized control layer—enabling forecast-based discharge into the energy community when demand and pricing conditions are favorable—can yield additional savings of €28 to €139 per year. These gains result from improved compensation for energy shared within the community and underscore the potential of coordinated storage management. In practice, the net benefit of this additional optimization is likely to be more modest, with a conservative estimate placing it between €10 and €60 per year due to forecasting uncertainty and operational variability.

Looking ahead, the financial outcomes for prosumers in energy communities could improve under more dynamic electricity pricing conditions. Time-variable pricing models, which are being gradually adopted by utilities and consumers, may offer additional opportunities for cost optimization—particularly as the number of hours with very low or even negative electricity prices increases, often during periods of high solar generation in summer months. In this context, adapting internal pricing structures within energy communities—for instance, by introducing incentives for feeding electricity into the grid during periods of lower solar production, such as early mornings or evenings—could encourage behaviors that support both individual savings and greater energy self-sufficiency at the community level. More broadly, aligning local energy use with market signals may enhance the overall economic viability of energy communities in the evolving energy landscape.

The findings support the view that intelligent, community-oriented storage operation can become a central component of future energy communities. Predictive tools, combined with coordinated dispatch strategies, offer a scalable path toward greater autonomy and more efficient use of local renewable energy.

Additional methodological details, household-level results, and scenario comparisons can be found in Appendix C (Section 9.3). This supplementary material provides further context for the findings presented here and illustrates how the reported outcomes were derived.

9.3 Appendix C: Supplemental Information Simulations

This appendix provides supporting information and methodological context for the findings presented in Section 5.2.3. While Section 5.2.3 summarized the key results of a simulated storage deployment in an energy community setting, the following sections detail the underlying assumptions, data sources, and modelling approach used to derive those results. Additional examples and disaggregated outcomes are included to illustrate how specific conclusions—such as the benefits of community-oriented storage operation—were reached. This supplementary analysis also includes further scenario comparisons and household-level insights that reinforce the observed trends and help clarify the potential and limitations of storage-based optimization in real-world energy communities. Illustrations courtesy of 4ward Energy Research GmbH (4ward).

9.3.1 Simulation Methods: Battery Storage

In contrast to the forecast-based simulations used in Section 5.2.3 to compare different forecasting approaches, this analysis focused on assessing the potential contribution of residential storage systems to increasing self-consumption within energy communities. It is based on historical data and therefore assumes a perfect forecast of both generation and consumption, allowing for a clearer comparison of storage operation strategies under idealized conditions.

Two energy communities provided by project partner neoom were selected to reflect differing structural and consumption characteristics. Within each community, two to three households equipped with photovoltaic (PV) systems—but without battery storage—were selected for the simulation. Using actual historical load and generation data, each household's energy profile was modelled, and the impact of adding a 20-kWh battery storage system was analyzed through scenario-based simulation.

Three operational scenarios were modelled to explore alternative approaches to storage use. A representative tariff structure was applied to evaluate economic impacts and assess potential cost savings under each scenario. Table VIII-Table X provide an overview of the tested scenarios, the tariff model used, and the characteristics of the participating prosumers.

Table VIII: Simulated Storage Operation Scenarios - Overview of Simulated Storage Operation Scenarios. Description of the three scenarios used to evaluate the impact of residential storage systems on energy distribution and economic outcomes within energy communities (4ward).

Scenario SN0	Base	Reflects the actual historical state of energy distribution without any storage system in place. This scenario serves as a reference point for evaluating the impact of storage-based optimization strategies.
Scenario SN1	Self-sufficiency optimization	Storage is charged when surplus PV generation is available and discharged when household consumption exceeds generation. Decisions are based purely on energy balance at the individual household level, with the aim of maximizing local self-consumption and minimizing reliance on grid supply.
Scenario SN2	Self-sufficiency + community optimization	Builds on the self-sufficiency approach but incorporates economic optimization by considering the higher compensation rate available for energy fed into the energy community. Battery discharge decisions are informed by day-ahead forecasts and are not limited to meeting household demand—energy may also be discharged strategically to supply the community when demand exists, thereby increasing individual revenue.

Table IX: Tariff Structure Used in Simulation - Tariff Structure Applied in Simulation. Electricity pricing used in the analysis. Consumption Grid and Consumption EC refer to the costs paid by end users for electricity from the public grid and the energy community, respectively, including both energy and grid charges. Feed-in EC and Feed-in Grid indicate payments received by prosumers for energy sold to the community or external supplier, based on the energy component only (4ward).

Consumption Grid (Energy + Grid)	29.60 ct/kWh
Consumption EC (Energy + Grid)	18.2 ct/kWh
Feed in Tariff EC	10 ct/kWh
Feed in Tariff Grid	6 ct/kWh

Table X: Characteristics of Simulated Prosumers - Characteristics of Simulated Prosumers. Overview of the households included in the simulation, including photovoltaic (PV) peak capacity and relevant property features that influence electricity demand, such as electric vehicles, heat pumps, and household size (4ward).

Ref.	PV-Peak Power	Properties
Prosumer A	7.77 kWp	Electric car and Heat pump
Prosumer B	27 kWp	n/a
Prosumer C	10 kWp	4 People, Heat pump
Prosumer D	10 kWp	Electric car and Heat pump, 4 Persons
Prosumer E	12.9 kWp	Heat pump, Wellness (Sauna/Pool)

9.3.2 Results: Simulated Storage

The results show a consistent pattern across all analyzed prosumers. When storage is operated solely for self-sufficiency (Scenario SN1), the amount of energy fed into the community decreases by 40% to 79% compared to the baseline without storage (Scenario SN0). This reflects the fact that surplus energy is retained for individual use rather than shared within the energy community. In contrast, under the combined self-sufficiency and community-optimized strategy (Scenario SN2), the volume of energy delivered to the community increases substantially—by 149% to over 1000%, also relative to the baseline—as storage is used not only to meet household demand but also to supply other community members when local demand exists and compensation is more favorable. This shift supports greater local utilization of renewable energy and reduces reliance on external energy suppliers.

Figure 31 compares two battery storage operation strategies across five prosumers, highlighting their impact on electricity flows and grid interaction. Grid Consumption refers to the reduction in electricity drawn from the public grid. Infeed COOP (rel) and Infeed COOP (abs) represent the percentage and absolute changes, respectively, in energy delivered by the prosumer to the energy community. The term COOP refers to the optimization platform used to manage intra-community energy exchange and is used here as shorthand for energy community interactions. Infeed EVU indicates the change in electricity exported to the external energy supplier. All values are shown relative to Scenario SN0 (baseline without storage).

It should be noted that extremely high percentage increases (such as the 1000%+ range) can occur when baseline values are very low. In such cases, even moderate absolute changes can yield large relative differences. These figures should therefore be interpreted in the context of the actual energy volumes involved.

		Gird Consumption	Infeed Coop (rel)	Infeed Coop (abs)	Infeed EVU
Self sufficiency optimisation	Prosumer A	-42%	-79%	-385,6 kWh	-44%
	Prosumer B	-91%	-40%	-918,3 kWh	-6%
	Prosumer C	-71%	-64%	-329,3 kWh	-35%
	Prosumer D	-73%	-61%	-127,7 kWh	-31%
	Prosumer E	-80%	-58%	-177,5 kWh	-22%
		Gird Consumption	Infeed Coop (rel)	Infeed Coop (abs)	Infeed EVU
Self sufficiency optimisation + **EEG optimisation**	Prosumer A	-38%	293%	1816,6 kWh	-78%
	Prosumer B	-84%	149%	4362,3 kWh	-28%
	Prosumer C	-65%	344%	2088 kWh	-63%
	Prosumer D	-65%	1051%	2314 kWh	-31%
	Prosumer E	-72%	787%	2599,4 kWh	-61%

Figure 31: Survey: Community Benefit - Comparison of storage strategies showing changes in grid consumption (decreasing), energy delivered to the energy community (Infeed COOP, increasing), and energy exported to the external supplier (Infeed EVU, decreasing). COOP refers to the community optimization platform used to manage intra-community energy exchange. All values are shown relative to Scenario SN0 (baseline without storage).(4ward).

These operational differences also result in measurable economic impacts. Under the self-sufficiency strategy, annual cost savings for individual prosumers range from approximately €360 to €610. By enabling additional energy sales within the community in Scenario SN2, further annual savings of €28 to €139 are achieved, depending on each household's consumption and generation profile. These additional benefits result from the higher feed-in tariff available for energy delivered to the community compared to the public grid.

It should also be noted that these figures reflect savings based on energy tariffs alone. The costs associated with the purchase, installation, and operation of the battery storage systems were not included in this analysis. As such, a portion of the reported savings would need to be allocated toward financing the storage investment in a full economic assessment.

Figure 32 presents the cost savings achieved by individual prosumers under the two storage operation strategies. The first column, Self-sufficiency opt, shows the reduction in electricity costs when battery storage is operated to maximize local self-consumption. The column Additional Full Opt represents the further annual savings generated through the community-optimized strategy, which allows stored energy to be supplied to the energy community when local demand exists and tariffs are more favorable. These additional earnings reflect the improved compensation available within the community compared to feed-in to the public grid.

		Self sufficiency opt	Additional Full Opt
Cost reduction	Prosumer A	-€ 559,21	-€ 28,09
	Prosumer B	-€ 468,52	-€ 138,99
	Prosumer C	-€ 607,01	-€ 32,43
	Prosumer D	-€ 474,35	-€ 37,50
	Prosumer E	-€ 364,72	-€ 68,46

Figure 32: Annual Cost Reductions - Annual cost reductions per prosumer under (1) self-sufficiency optimization and (2) full community-aware optimization. "Additional Full Opt" reflects the extra savings gained by supplying energy to the energy community at higher feed-in tariffs (4ward).

The additional value of community-oriented storage optimization can be illustrated more concretely by looking at the results for an individual prosumer—in this case, Prosumer C. Figure 33 compares the three scenarios and highlights how energy flows shift under each strategy. In Scenario SN2, the volume of energy delivered to the energy community increases by approximately 1,500 kWh. At the same time, grid consumption is reduced by 2,500 kWh, and feed-in to the external supplier (EVU) is reduced by 4,000 kWh. This outcome demonstrates the combined effect of optimized battery operation and forecast-based scheduling in increasing local energy use and improving economic returns.

Prosumer C represents one of the more favorable cases in the analysis. This may be due to a household load profile that includes energy consumption during early morning or evening hours, which can be covered by stored solar energy. In addition, surplus PV generation during the day can be sold into the energy community when local demand exists and pricing conditions are more advantageous. These factors reinforce the value of predictive, community-oriented storage strategies—though results will vary depending on each household's specific generation and consumption patterns.

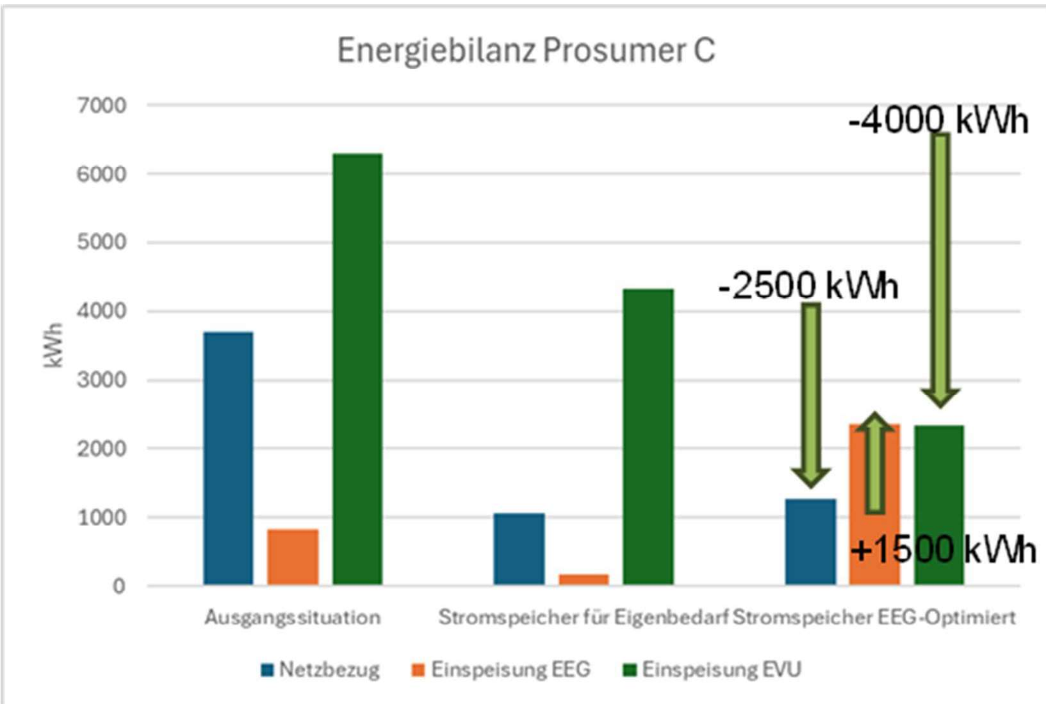

Figure 33: Scenario Comparison: Prosumer C - Energy distribution comparison across scenarios for Prosumer C. Ausgangssituation (EN: SN0 Baseline scenario), Stromspeicher für Eigenbedarf (EN: SN1 Storage for self-consumption), and Stromspeicher EEG-Optimiert (EN: SN2 Storage with community optimization). The bars represent Netzbezug (EN: Grid consumption), Einspeisung EEG (EN: Feed-in to energy community), and Einspeisung EVU (EN: Feed-in to external supplier).(4ward).

9.4 Appendix D: Data Validation and Preprocessing

As outlined in Section 6.2: *Monitoring and Analysis Methods*, a key part of the project's evaluation approach involved analyzing real-world energy data from the testbeds to assess the performance of different energy community configurations. Reliable input data was essential for calculating energy indicators, simulating alternative scenarios, and estimating potential benefits such as self-sufficiency and cost savings. This appendix provides a detailed overview of the data preparation process used to clean, validate, and complete the measurement data before analysis.

The measurement data provided by project partners—including outputs from photovoltaic (PV) systems, battery storage, and the power grid—were analyzed and prepared to establish a reliable foundation for evaluating energy flows and community configurations. The objective was to identify anomalies, address missing values, and ensure that the data were both complete and physically plausible.

In addition to cleaning and validating the data, various energy indicators were calculated, including self-sufficiency and self-consumption ratios. These metrics informed the modeling of different energy community scenarios, using both static and dynamic allocation methods. A profitability analysis was also conducted to estimate potential economic benefits and identify areas for optimization.

The data preparation process involved four main steps:

Step 1: Structuring and Initial Validation

Measurement data were extracted from provided CSV files and standardized into a consistent format. Initial plausibility checks were performed, including a review of the five highest PV production values per site. These peak values were later cross-checked to confirm consistency with system size, helping to detect and correct any unrealistic values.

Step 2: Addressing Missing Battery Storage Data

Missing battery values were completed using simplified assumptions. In cases of brief data outages, the battery was assumed to act as a passive energy conduit—passing energy from the PV system to the building or grid without charging or discharging. This approach avoided the introduction of additional errors from simulated battery state-of-charge (SOC) tracking.

Step 3: Filling PV Gaps and Removing Outliers

PV production gaps were filled by referencing a comparable system in the same municipality with similar orientation. The reference data were scaled according to system size to reconstruct missing values.

Outlier filtering included:

- Removing values exceeding 110% of the installed PV system's capacity.
- Deleting small PV readings during nighttime hours (10 PM to 3 AM).
- Ensuring that calculated energy consumption values were never negative—any such cases were corrected by adjusting PV production to restore energy balance.

These adjustments addressed common errors such as inverted meter readings or incomplete substitutions from reference data.

Step 4: Completing Grid Power Data and Calculating Consumption

To complete the dataset, remaining gaps in grid power values were estimated using neighboring time points. Where historical consumption values existed one or two weeks before or after the missing timestamp, they were used to interpolate or average expected consumption. Grid power was then back-calculated using the energy balance equation (consumption = PV + battery + grid). This method ensured that all energy flows could be reconstructed as accurately and consistently as possible.

The resulting dataset enabled reliable analysis of energy flows between PV systems, batteries, and the grid. The combination of plausibility checks, rule-based substitution, and historical interpolation allowed for a fully automated and repeatable data preparation process.