Musterdeckblatt für Abschlussarbeiten

Bachelorarbeit / Masterarbeit
Zur Erlangung des akademischen Grades Bachelor ... /Master ...

Master of Science

Thema der Arbeit

Genetic Algorithm Optimization of Charging Infrastructure Locations
for the Operation of Battery-Electric Trains in German Regional Passenger Rail

eingereicht von: Name der Verfasserin / des Verfassers

Emil Carlé Dietrich

Gutachter/innen: Titel, Name

Titel, Name

Prof. Dr. Tobia Lakes

Dr. Stephan Schmid

Eingereicht am Geographischen Institut der Humboldt-Universität zu Berlin

Genetic Algorithm Optimization of Charging Infrastructure Locations Enabling the Operation of Battery-Electric Trains in German Regional Passenger Rail

Master's Thesis by Emil Dietrich

Geographisches Institut

Humboldt Universität zu Berlin

05.10.2023

Abstract

Responding to global climate change and political aims at the European level, the German government aims to achieve carbon neutrality by 2045. Embedded in this aim lies a necessity for converting diesel multiple units to electric multiple units – especially on regional rail passenger routes where diesel multiple units are often used. Making great use of existing electrification and longer layover times, battery-electric multiple units are particularly well-suited for these routes. However, identifying the best the spatial allocation of charging infrastructure to allow for their operation constitutes a complex non-linear optimization problem. This thesis develops a standardized methodological concept based in a genetic algorithm to optimize this spatial distribution with the least infrastructure cost using vehicle circulation plans, simulated vehicle power at wheel, OpenStreetMap data, a digital terrain model, a timetable dataset, and route geometries. The approach is developed and tested on a circulation plan for the German subnetwork of Pfalznetz. The methodology successfully identifies optimal spatial distributions of charging infrastructure but suffers from shortcomings relating to the quality and availability of input data, the computational efficiency of the algorithm and the approaching of local minima. Specific minor alterations to the code may significantly improve the quality of the results. The flexibility and broad applicability of the method primes it for expansion accounting for a wider range of aspects relating to charging infrastructure placement.

Acknowledgements

I sincerely thank the de facto supervisors who assisted, discussed, and evaluated the thesis contents throughout its progress. I appreciate the help they provided and find myself thankful for the time they set aside to aid the thesis in its development.

De Facto Primary Supervisor

Sebastian Herwartz

Deutsches Zentrum für Luft- und Raumfahrt

De Facto Secondary Supervisor

Johannes Pagenkopf

Deutsches Zentrum für Luft- und Raumfahrt

De Facto Secondary Supervisor

Tobia Lakes

Humboldt Universität zu Berlin

Table of Contents

Table of Figures

Table of Tables

1. List of Abbreviations

BEMU	Battery-Electric Multiple Unit
DB	Deutsche Bahn AG
DEAP	Distributed Evolutionary Algorithms in Python
DMU	Diesel Multiple Unit
DLR	Deutsches Zentrum für Luft- und Raumfahrt (The German Aerospace Center)
DTM	Digital Terrain Model
CPU	Central Processing Unit
EMU	Electric Multiple Unit
FCEMU	Fuel-Cell Electric Multiple Unit
GTFS	General Transit Feed Specification
IPCC	Intergovernmental Panel on Climate Change
OSM	OpenStreetMap
SoE	State of Energy

2. Introduction

This section introduces the concept and overall research aim of the thesis together with a thematic introduction placing the thesis in a greater context. The section first highlights the direct connections between global climate change, political ambitions recognizing the need to act, and the necessary changes at the level of rail infrastructure needed to accommodate these changes. The section then elaborates on the research goals of the thesis and how the methodology and thesis content is built and structured around these aims.

2.1 Problem Statement

Global climate change is one of, if not the, greatest challenge that ever faced humankind. According to the IPCC human activities have already resulted in a global average surface temperature increase of 1.1 °C in 2011-2020 compared to 1850-1990 levels [1]. The wide impact of global climate change makes up one of the most pervasive forces driving global change in the 21st century. It is a meaningful threat to the sustainability of humankind, global well-being, and the ecosystems of planet Earth. This threat is anticipated to grow due to further incremental increases in average surface temperature from existing and future greenhouse gas emissions. Whether the warming can be limited to 1.5 °C or 2 °C targets depend on humankind's collective ability to reach net zero greenhouse gas emissions [1].

Responding to policies at the European level, the German government enacted a new Federal Climate Protection Act on in 2021 detailing and updating targets and strategies relating to German greenhouse gas emission reductions [2]. The government aims to reduce national greenhouse gas emissions by 55% by 2030 and 88% by 2040 compared to 1990 levels [2]. The political aim is to reach carbon neutrality by 2045 [2].

In an aim to reduce greenhouse gas emissions, decarbonization of energy consumption requires systemic transitions in all sectors of society. This is especially true for the transportation sector where vehicles often operate on diesel. Here, changes to the grid energy mix cannot aid in decarbonizing. Instead, systemic changes need to be implemented to allow for operation of different vehicles relying on non-fossil fuels. In Germany, the transport sector accounts for 20% of CO_2-equivalent emissions [3]. These emissions include those from rail although this is not the main contributor to the figure. However, focusing on the railway sector remains crucial to achieve net neutrality – especially in a world where public transit is increasingly necessary as an alternative to the intensive energy demands of private transportation.

In rail infrastructure, electrified networks have long existed alongside those operated by fossil-fuels. Due to higher operation cost efficiency, electrified rail infrastructure is often preferred for train routes through denser areas with high energy demand from frequent acceleration, or on high-speed long-distance routes requiring larger amounts of energy to maintain speed. DMUs are preferred for regional, less frequent routes with larger distances between stations. This is often due to the high upfront cost of catenary infrastructure not matching commuter demand or where catenary is not feasible in tunnels, mountainous terrain, or other obstacles. However, to decarbonize the railway system, DMUs must increasingly be replaced by alternative vehicles running on low- or zero tailpipe emission technology - even on routes where DMUs are traditionally chosen as the superior option. While DMUs do also operate elsewhere, regional commuter rail is currently predominantly serviced by DMU vehicles.

The railway infrastructure of Germany is comprised of a total of 33286-39773 km of track in 2020 [4], [5] (depending on the source). Of this, 20365-21100 km (53%-61.2%) are currently electrified [4], [5] (depending on the source). To replace all DMUs with EMUs that operate with constant access to the energy grid, large and costly infrastructure investments are required to install catenary systems on the remaining non-electrified track. While technically feasible, the cost of the endeavor encourages the examination of alternatives.

Often discussed alternatives to DMUs outside multiple units relying on catenary includes FCEMUs and BEMUs that both allow for a transition away from DMUs without the installation of catenary (or other electrification infrastructure) on all railway tracks. With an ability to operate under conditions of partly electrified railway track without the construction of new infrastructure projects, BEMUs offer a distinct advantage over other DMU alternatives. BEMUs are capable of using existing electrification infrastructure where it is already present and are particularly suitable for regional routes where layover times are often longer. This opens more avenues for cost-efficient charging in terms of infrastructural requirements. However, the feasibility and scope of optimal replacement of DMUs with BEMUs is a complex problem depending on the structure and existing electrification of the railway network, train schedules, and train energy consumption. It is the solving of this particular problem that constitutes the main aim of the thesis.

2.2 Research Aims

This thesis presents a methodological concept to identify the lowest-cost spatial allocation of charging infrastructure allowing for the continued operation of a passenger rail vehicle circulation plan after converting all vehicles to BEMUs. The optimization problem is spatial. Charging infrastructure is more efficient in one location than another, and identifying the most optimal solution is a complex non-linear optimization problem in space. It is for this reason the field of geoinformatics is of particular relevance to the problem area.

At its core, the thesis is a methodological exploration. The core goal is to develop a functional standardized methodology capable of identifying the optimal spatial allocations using a generalized set of input data. I present a methodology with this aim based on the application of a genetic algorithm. The approach is structured from end-to-end as a set of integrated Python scripts using a vehicle circulation plans, simulated vehicle energy demands, a GTFS dataset, OSM data, and a DTM to output a spatial allocation of BEMU charging infrastructure. The core research question can be summarized as:

How can a genetic algorithm be used to optimize the spatial allocation of charging infrastructure allowing for the operation of BEMUs on a regional passenger rail network with the lowest infrastructure cost?

The optimization problem is complex due to both its combinatorial nature and the difficulty with which to implement optimization functionality while reasonably asserting a real-world situation. The genetic algorithm is selected due to its flexibility in application and ability to handle combinatorial problems. The presented methodology accounts for existing electrification, handles parallel and single tracks, can accommodate any number of vehicles circulating simultaneously on the same or different tracks, accounts for layover times between individual trips, and understands infrastructure costs as both by electrified track length and by independent contiguous islands of catenary.

To ascertain the performance of the methodology, the thesis is built around a vehicle circulation plan for the German regional passenger rail subnetwork of Pfalznetz. Whilst the thesis explores in depth the results of the methodology as applied to Pfalznetz, the study of Pfalznetz is not the goal of the thesis itself. The dataset is used both to compare the results to an existing BEMU electrification plan for the region and to perform a sensitivity analysis of the applied method. This can be reduced to the following key questions:

What is the optimal spatial allocation of charging infrastructure allowing for the operation of BEMUs in the regional network of Pfalznetz?

To what extent does the implemented optimization approach successfully optimize spatial allocation of charging infrastructure for BEMU regional passenger rail?

3. State of the Art

This section provides an overview of the state of the art in scientific literature relating to the current state of battery-electric train technology, approaches to optimization of charging infrastructure in various network scenarios, and general applications of the genetic algorithm as an optimization tool in complex non-linear combinatorial problem-solving. Synthesizing these three core areas of research, this section attempts to situate this thesis in the greater scientific literature existing at the intersection between advances in train engineering, spatial analysis of networks, and metaheuristic algorithm optimization.

3.1 Sustainable Mobility

Literature on the concept of sustainable mobility is exploring the meaning, feasibility, and goals of the future mobility system trying to examine the theoretical direction for the transport of tomorrow. This study embeds itself within, and positions itself in relation to, the environmental aspects of the sustainable mobility concept. Although acutely relevant for the actual realization of truly sustainable mobility systems, questions on social or economic sustainability are not discussed in this thesis.

E. Holden et al. presents a convincing conceptual overview of the narratives contained within the idea of sustainable mobility [6]. Centrally, the paper argues that sustainable mobility is usually conceptualized (in politics and in research) in one of three overarching ways: electromobility, collective transport 2.0, and low-mobility societies [6]. The three grand narratives are distinguished by how they envision sustainable mobility to be realized. The electromobility narrative argues for the electrification (or 'hydrogenification') of the mobility system to achieve zero environmental impact, the collective transport 2.0 narrative argues that we must further collectivize and consolidate movement to increase overall efficiency, and, lastly, the low-mobility society narrative argues that the reduction of movement altogether is the key solution to a sustainable mobility system [6].

In order to place this thesis within a greater theoretical context, it is worth noting its strong lineation towards the electromobility narrative. The simple idea that a conversion from diesel- to electric- operated railway constitutes the core element of the transition towards environmentally sustainable mobility underpins the motivation of the research. While I maintain the opinion that the transition away from fossil fuels is a key element of a sustainable mobility system, I recognize the limitations of this narrative as the sole means of achieving sustainable mobility. As pointed out by E. Holden, even a complete electromobility transition still requires resource intensive means of production to keep it operating and expanding [6]. This has an inevitable environmental impact. Arguably, a holistic synthesis of all three grand narratives poses the most convincing understanding of the future sustainable mobility system.

The few studies that have attempted to evaluate the best pathway for a sustainable electromobility transition have not achieved any clear consensus on the direction that either should or will be followed. An analysis of green mobility (environmentally focused sustainable mobility) strategies for the European mobility system concludes that complete electrification of rail and road mobility is the most cost-efficient pathway for decarbonization of the European mobility system [7]. This is done comparing different vehicle composition scenarios. Germany, one of the three countries used in the cited case study, is also evaluated to be the most cost-efficient in investing in electrification infrastructure despite higher estimated energy costs. In the literature search done within this study, I have not found any conclusive studies combining concepts of sustainable mobility with strategic consensus on how to best proceed with the systemic transition.

3.2 Battery-Electric Train Technology

With recent advances in electric battery technology allowing for the flexible electrification of infrastructure that was once bound to fossil fuels, the forefront of engineering of BEMUs provides new opportunities for the future of rail infrastructure. Like the development of battery-electric cars, research in, and development of, battery-electric train infrastructure is currently experiencing rapid advances. This section briefly explores a small set of literature examining the economic and environmental potential of BEMUs. This section also presents single key studies from which crucial assumptions and simplified understandings of BEMU charging infrastructure is drawn. In keeping with the geoinformatics and genetic programming focus of this thesis, the engineering aspect of the literature review is kept minimal.

Given the recent, and continuous, emergence of BEMUs as viable alternatives to existing DMUs and other green transition options like FCEMUs, some studies have focused on evaluating their feasibility through various economic and environmental assessments. Popovich, N. D. et al. performs a comprehensive combined cost and environmental impact assessment for converting DMUs in the U.S. to BEMUs [8]. The study found that a 241-km driving range could be achieved for BEMUs replacing the existing DMUs, and that these would be able to operate while consuming 50% less energy [8]. This energy could also be electric, allowing for a green energy supply. The study concludes that that the U.S. freight sector could reduce costs by $94 billion over 20 years when accounting for the costs of air pollution and CO_2 emissions within the cost model. Other studies, like Mwambeleko, J. J. et al. examining a 12 km commuter route in Tanzania, have studied the viability of replacing DMUs with BEMUs on specific rail routes [9]. The study found that the replacement of the current operating DMUs with BEMUs could reduce the fuel costs by 79% [9].

The charging infrastructure required for the operation of BEMUs are a significant part of the costs associated with moving from DMUs to BEMUs. The techno-economic assessment by Streuling, C. et al. [10] of BEMUs and charging infrastructure forms the basis for how this infrastructure is conceptualized (and simplified) within this thesis. In the study, the researchers analyze the potential for cost reduction of charging infrastructure comparing different configurations of electrification infrastructure alternatives. From this study, I draw the simplified understanding of electrification infrastructure as composed of overhead catenary and substations and their respective capital and life cycle costs. In order for a vehicle to charge, it must be on track with installed catenary. In line with the assessments with the study, I assume a catenary cost of €610/m [10] . In addition to this, all independent islands of catenary, referring to contiguous stretches of catenary disconnected from other catenary islands, must be served by a substation converting voltage and frequency from either the public or the railway electricity grid. Refer to Figure 8 in section 4.5.4 for a visualization of the concept of catenary islands. Aligning with the study by Streuling C. et al, I assume a substation cost of €8 million [10]. While this particular study distinguishes between these two infrastructures, I do not evaluate energy draw from the perspective of the grid and consider both substations identical [10]. The existing catenary is already supplied with a substation and does not need additional substations if it is extended. This may not be accurate if the capacity of the existing substation cannot service the new demand, but such a problem is not evaluated within the scope of the thesis.

3.3 Applications of the Genetic Algorithm

Advances in computation power continue to enable new algorithmic approaches to problem solving that were once barred behind computational limitations. The genetic algorithm, once nearly impossible to execute with available computational technology, has recently seen a sharp rise in use due to its conceptual simplicity and flexibility in application. The algorithm is specifically suited for complex optimization tasks of problems of a non-linear and combinatorial nature, and the algorithm is quickly gaining adoption within research dealing with such problems. This section introduces the core workings of the genetic algorithm before exploring its wide range of applications in modern optimization problem-solving.

3.3.1 The Genetic Algorithm: An Introduction

A genetic algorithm is a metaheuristic algorithm attempting to identify the best solution in a problem space by creating and evaluating a series of possible problem solutions. This is distinctly different from mathematical optimization along a discrete or continuous space where the optimal solution is identified on a pre-determined axis. Metaheuristic algorithms are especially advantageous over standard mathematical optimization when the optimization problem is either combinatorial or non-linear. Depending on the problem at hand, non-linear optimization problems can be expressed in linear space or otherwise altered in order to fit mathematical optimization approaches, but this is not always feasible and can require layers of abstraction that compromise the complexity of the problem addressed. Metaheuristic algorithms do not rely on pre-determined axes for optimization and therefore are indifferent to the linearity, or non-linearity, of the problem space. Combinatorial problems are especially difficult to align with mathematical optimization because no incremental change in any value is sufficient to approach an optimum. Metaheuristic algorithms circumvent this problem by allowing for custom evaluations of solution suitability that depend entirely on the complete composition of inputs in the problem space and not on any single value. With recent advances in computational power, the high computational requirement of generating and evaluating series of solutions have started to pave the way for metaheuristic algorithms as commonplace in optimization problem-solving [11].

Metaheuristic algorithms, such as particle swarm optimization, ant colony optimization, or the genetic algorithm, draw inspiration from biological processes [11]. The genetic algorithm, initially conceptualized and developed throughout the 1990s, draws inspiration from basic evolutionary principles [11], [12]. The metaheuristic process of the algorithm expresses each generated possible problem solution through an encoding scheme akin to that of a chromosome. The most commonly applied encoding scheme - which is also employed in the research of this thesis is the binary encoding where the problem solutions are expressed as strings of binary bits [11], [13]. Other encoding schemes include octal and hexadecimal, where the solutions are encoded octal or hexadecimal numbers, and value encoding, where each number in the string refers to the value of a specific variable [11], [13].

Genetic algorithms can be conceptualized with basic evolutionary and biological principles. The algorithm initiates by generating a set of possible problem solutions. Each possible problem solution, referred to as an individual, is expressed through the encoding scheme. The individual can be thought of as a genome, and the encoding scheme as its genes. The entire set of generated individuals is referred to as the population. In biology, this can be imagined as a group of organisms each with its own genome. Each individual is evaluated on their ability to solve the target problem. The evaluation process is defined through a fitness function that assigns a fitness value to each individual representing how well the individual solves the problem. Within Darwinian evolution, this refers to each organism's ability to survive. The internal workings of fitness function can take any form and is entirely dependent on the encoding scheme of the individuals and how this information relates to the problem the generic algorithm is solving. Based on the fitness value of each individual in the population, some are selected for further processing through the selection technique. The selection technique describes the way in which individuals are selected from the population by fitness value through a semi-stochastic process (depending on technique) [11], [14]. This can be understood as a process of combat, for examples when males in a flock fight each other for the ability to mate. Common approaches include tournament, wheel selection, and stochastic universal sampling. After selection, the encoding of the select individuals is used in the crossover operator to populate a new generation of individuals. In biology, this process is akin to that of mating. Crossover operators, like single-point, k-point, and shuffle, all mix and match parts of the individuals in order to create offspring carrying qualities of the fittest individuals from the prior generation [15]. Finally, some values in some individuals in the new population are altered according to the mutation operator. Mutation, as the name suggests, follows the same concepts as mutation of genes within biology. Mutation operators, like simple inversion and scramble mutation, attempt to prevent convergence on local optima and allow the algorithm to explore more of the problem space [11], [14]. The genetic algorithm continues to iterate through generations of populations until a termination criterion, specifying when an optimal solution is found, is satisfied (refer to Figure 1).

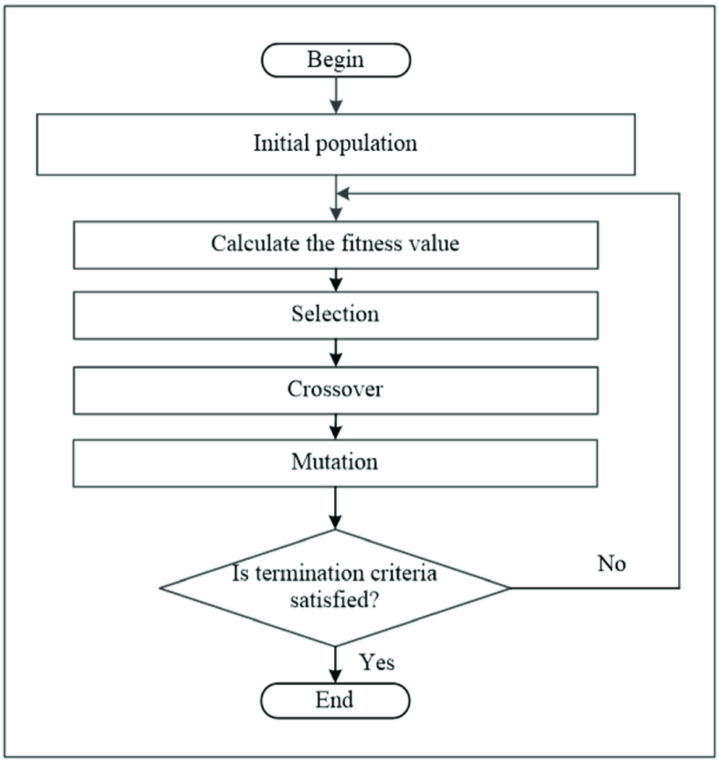

Figure 1: Conceptual illustration of a genetic algorithm [16].

3.3.2 The Genetic Algorithm as an Optimization Tool

Variations of the genetic algorithm are applied in optimization tasks across disciplines. It has been used in fields ranging from game design and medical research to image processing and precision agriculture [11], [17], [18]. Especially relating to the general field of geography, its applications in image processing and precision agriculture highlight the versatility of the optimization tool. This versatility also aligns the algorithm well with tasks of network optimization.

In image processing, use of the genetic algorithm has aided in the optimization of image segmentation, denoising, object detection, and object recognition [11], [19]. An exploration of the genetic algorithm's performance in image segmentation tasks in comparison to other soft computing approaches like fuzzy logic and artificial neural network concludes, in line with other literature, that the genetic algorithm is a suitable candidate in the development of image segmentation techniques [20]. For example, developing further on clustering the validity indexing techniques that are used in image segmentation, genetic algorithms have been found to be helpful in improving the accuracy of super pixel clustering [21]. The genetic algorithm has also often been used to reduce image salt and pepper noise [19]. Using an iterative approach through a genetic algorithm instead of a standard filter, the genetic algorithm approach was found to outperform other common methods of reducing salt and pepper noise in images [22]. The continued development of image processing techniques, be it through genetic algorithms or otherwise, percolate into several research applications.

Precision agriculture is one such application. Specifically due to the genetic algorithm's application in image segmentation and its ability to work with complex combinatorial optimization problems, some studies have used it for weed detection in agricultural settings [23]. In one such study, the genetic algorithm was applied to search for optimal light conditions to best identify weeds across images with varying light intensities. The supervised image segmentation was significantly augmented with the integration of the genetic algorithm approach. Another study successfully employed a genetic algorithm to optimize the distribution of nitrogen fertilizer on a field balancing the target nitrogen application per field grid cell with equipment strain resulting from sharp sudden shifts in application amounts [24]. Lastly, examining more fine-scale data, a study employed a genetic algorithm to extract and identify taxonomic features and species identification from leaf fragments of young sparse-leaved canopies growth in greenhouse and field conditions [25]. The algorithm was able to identify at least one plant leaf from imagery of the entire plant in 92% of cases after three weeks of plant growth [25].

3.4 Spatial Optimization in Transit Networks

Spatial optimizations of network configurations continue to present challenging obstacles to overcome for researchers regardless of the exact thematic nature of the problem. In solving spatial network optimizations, researchers have employed a large variety of optimization approaches and algorithms depending on the problem specifications and data availability. This section explores highlights within spatial network optimization literature before homing in on the relatively sparse research on optimization specifically relating to rail network charging infrastructure. This section also briefly examines the use of GTFS data in analysis of public transit networks.

3.4.1 Use of General Transit Specification Feeds

Originally developed by Google in 2015, the GTFS format has since gained dominance as the de facto standard format for storing public transit timetables [26], [27]. The continuous adoption of the data format by transit agencies combined with an increased quantity of freely released GTFS datasets has allowed researchers a new approach to studies dealing with public transit infrastructure. Building on the growing mobility research using GTFS data, this study chooses to use GTFS data on the German rail network as a cornerstone of both the data source for understanding the operation of the rail network and the Python program structure.

The location of transit lines is generally not difficult to access. What the GTFS dataset offers is information on when and where each stop is serviced by a transit vehicle. Many studies have attempted to use the information to create more precise evaluations on accessibility. Understanding a network as a collection of trips instead of a collection of lines and edges, one study used GTFS data to propose a new way to construct networks for transit system modelling [28]. Creating a network from GTFS data allows for the integration of temporality - sometimes a route is available, sometimes it is not. This enhances the possibilities and accuracies of accessibility studies involving public transit. This potential is tapped into by the public transit arrival time prediction models developed based on GTFS datasets in a study comparing the feasibility of different machine learning approaches [29]. This study found that combining neural network models, rigorous approaches to data validation, and GTFS datasets could together successfully be used to predict arrival times of public transit vehicles under different conditions [29].

Still not very widely applied, much research in the field still focusses on expanding the use cases of the GTFS datasets rather than directly generating new knowledge from available data. Arguably, much of the work of this thesis also falls within this line of research. The R-package gtfs2gps is developed by a group of researchers to allow for the conversation of GTFS dataset format to moment-to-moment GPS locations of transit vehicles [30]. This sizably expands the potential of use for GTFS dataset by further research but does so introducing uncertainty that should be discounted as GTFS data is abstracted to more generalized use. Some studies have worked on evaluating the limits of GTFS data for these kinds of research applications. Taking a critical look at GTFS dataset's' ability to be used to accurately estimate travel times and accessibility measures, one study found that predictions based on GTFS data over- and underpredicted travel times by up to 15% compared to known vehicle locations [31]. The study highlights the important fact that (static) GTFS data does not allow for the actual real-time tracking of vehicles and using the data as such comes with a sizeable margin of error.

3.4.2 Optimizing the Spatial Allocation of Charging Infrastructure

The solving of spatial allocation problems, such as the identification of optimal placement of charging infrastructure in a BEMU operated rail network, is a common challenge in network and transit infrastructure planning. The optimal placement of stations, transit lines, bus routes, road network configurations, etc. all serve as key areas of investigation for good infrastructure planning. Focusing specifically on charging infrastructure, a large quantity of studies has developed methodologies for optimal spatial allocation be it for cars, buses, or rail.

The spatial allocation challenges for rail infrastructure are not much different than those for buses and cars. Some studies have employed similar optimization approaches to that of this thesis while some have attempted different takes on the optimization challenge. Tzamakos, D. et. al [32], Wang, X. et. al [33], and Uslu, T. and O. Kaya [34] build their studies attempting to optimize charging infrastructure for electric buses on integer linear programming models. This is a common approach alongside evolutionary programming. Tzamakos D. et al. [32] optimizes bus charging infrastructure using a network analysis structure consisting of nodes and edges where the nodes are electrifiable. This setup generally makes the most sense for infrastructure development that is not placed along routes but at stops only. Unique to the setting of buses, this study and Uslu, T. and Kaya, O. [34] both attempt to implement a way in which to account for charge queueing, which is generally not a factor relevant to train infrastructure. Wang, X. et al [33] is built on a similar methodology but notably focuses on including an evaluation of optimal battery sizing within the optimization model itself. This way, the optimal spatial distribution of charging infrastructure is determined in unison with the relevant parameters of the vehicle for which the charging infrastructure must function.

In large part due to its versatility in application, the genetic algorithm is frequently used to optimize the placement of charging infrastructure within literature on both cars, buses, and rail. Arguably, this (and other evolutionary programming methods) may be more frequent than linear programming. Chen G. et al. has conducted studies similar in goal to those described above but using genetic algorithm approaches instead [35] [36]. For one study, the research group uses an interesting combination of methodologies applying a mixed integer non-linear model nested within a genetic algorithm allowing the genetic algorithm to tune the non-linear model [35]. This study was also built on routing information akin to that available in this thesis. The integration of this data extending into the charge waiting times trying to both optimize infrastructure costs while minimizing the impact on transit time for passengers. The other study by Chen et al., a case study in Xi'an, China, used a standard genetic algorithm approach also accounting for battery capacity similar to the studies mentioned above [36]. Additionally, this study also included evaluations on battery lifespan [36].

Electrical charging infrastructure for cars is different than that of buses due to their lack of consistent and predictable routing. Despite this, the methodologies for identifying optimal charging stations for cars are often similar in their core methodology. Variations of genetic algorithm approaches are not uncommon within the literature. Performing a case study in Milan, Akbari et al. [37] employs a genetic algorithm in order to identify the necessary number and spatial distribution of charging stations for cars. A challenging aspect of this methodological approach, a challenge that is shared in railway and bus systems alike, is the way in which the problem is codified for a genetic algorithm to optimize. For this study, demand and car movement was aggregated into areas, and these areas served as distinct elements for the algorithm to optimize. Expanding further on the use of genetic algorithms, it is also possible to nest a genetic algorithm within another to allow for programmed parameter optimization. This approach, a hierarchical genetic algorithm, is employed by Yan, X. et al. [38] to optimize charging infrastructure placement both accounting for infrastructure costs and constraints of the energy grid. The inclusion of the energy grid into the study is interesting in particular: since the studies relating to optimizing the locations of charging infrastructure operates with some understanding of the energy demand of the vehicles, the optimized result includes information on the energy draw at each identified charging location. It is therefore generally feasible to include the perspective of the energy grid relating to the power draw (refer to section 6 for a deeper discussion on this topic related to the methodology of this thesis).

Compared to buses and cars, fewer studies seem to have focused on rail infrastructure in the context of optimizing charging infrastructure. This may be due to the fact that electric car and bus vehicles have experienced a higher rate of market penetration and adoption than BEMUs. Bridging the gap between literature on buses and rail, Baumeister et al. [39] develops a method for optimizing catenary for a trolleybus system. In many regards, the methodology outlined herein shares similarity to that developed within this thesis. The data input used in the optimization approach is almost identical relying on simulated vehicle energy demand, timetables, and information on the road network. This model also attempts to account for existing infrastructure. The success of this model indicates that it may too be an applicable methodological concept for optimization on rail networks. This thesis attempts to expand on this part of the literature.

It is worth noting that the genetic algorithm is not the only approach within evolutionary programming that has been applied to the optimization of charging infrastructure placement. Looking specifically at railway networks, Roch-Dupré, T. D. et al. [40] compares different nature-inspired machine learning approaches (genetic algorithm, particle swarm optimization, and fireworks algorithm) on their optimization performance. However, while this study does deal with railway infrastructure, the key aim is to optimize the placement of energy storage systems and not to optimize the placement of charging infrastructure from the perspective of the vehicle. It is therefore difficult for the findings of the study to shed light on the best methodology for charging infrastructure placement.

4. Methodology

This section describes in detail the methodology of the thesis from initial input data and scope of research to the final output in the form of optimized spatial allocations of charging infrastructure for BEMU circulation. The method design presents the overall concept and aim of the methodology structure together with an overview of the way in which input data is formatted to allow for optimization based in a genetic algorithm approach. I then detail the choices related to the study area and scope followed by detailing the parameters of the generic BEMU used in the study. Each distinct input dataset is presented followed by a detailed chronological presentation of the data formatting process and the genetic algorithm implementation.

4.1 Method Design

The methodological concept of the thesis is to express the relevant railway network as distinct segments that can either be electrified or not and contain information on the BEMUs that traverse them and the theoretical battery SoE of the vehicles at the time of traversal. This segmented understanding of the network is expressed as an individual in the genetic algorithm. Each individual represents a possible configuration of additional electrification of the network (refer to Figure 7 in section 4.5.4). The genetic algorithm refines the individuals through its runtime before outputting a final "most-optimal" individual. The genetic algorithm is built to reduce the cost of additional electrification infrastructure while preventing any vehicle from dipping below a minimum battery SoE threshold throughout its circulation.

A schematic overview of the method design is visualized on Figure 2. Refer to relevant following subsections in section 4.5 on details on each aspect of the method. All input data is pre-processed before initiating the final data processing at each instance of a genetic algorithm run in the algorithm pipeline. The input vehicle circulation plan, the GTFS dataset, and the routed geometries are linked as a single the circulation plan (refer to Appendix A, Table 12). The routed geometries are enriched with OSM data and formatted as a single segmented network. Using a standardized set of input parameters for each genetic algorithm run the relevant network segments and vehicle energy demand simulations are selected at the start of each run. The simulations undergo relevant re-formatting and the potential change in battery SoE is calculated for each network segment the vehicle passes through. The trip simulations are concatenated for each vehicle in the selected parts of the circulation plan. The genetic algorithm optimizes the presence of new electrification on the network segments by iteratively altering the state of electrification at each segment. The algorithm is penalized for high electrification infrastructure costs and for not providing sufficient charging infrastructure to allow for the operation of the BEMUs above the assigned SoE threshold. The final output is the most optimal configuration of electrification infrastructure identified during the model run.

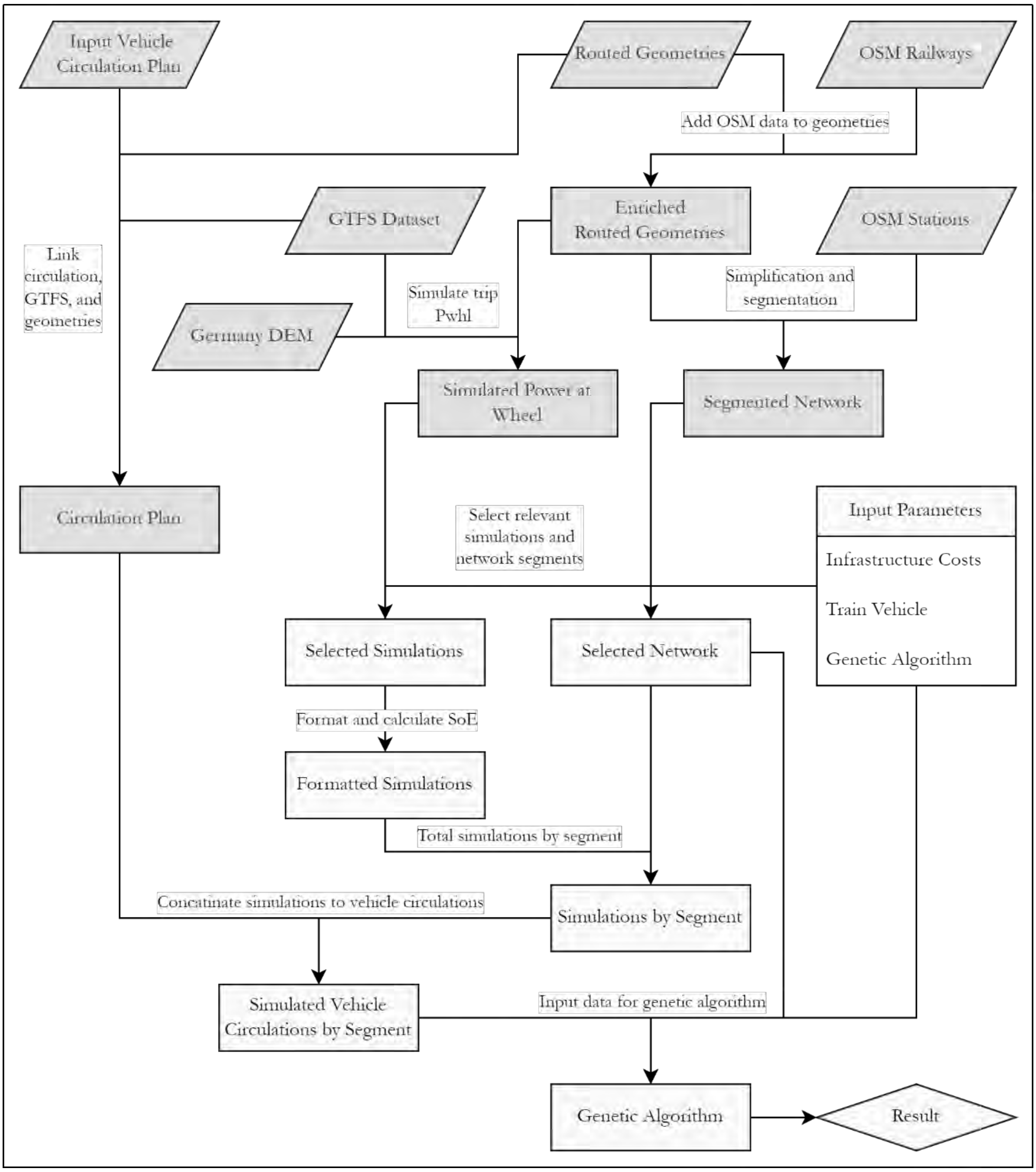

Figure 2: Schematic overview of the thesis methodology. Parallelograms indicate input data; rectangles indicate processed data. Preprocessed steps before genetic algorithm runs are grey while steps repeated at each genetic algorithm run are white. The result is visualized as a diamond.

4.2 Scope and Study Area

The approach for the optimization of charging infrastructure placement is designed to be applicable on any subset of a train network presuming the data is available. This is the case for the entirety of Germany for all relevant datasets besides the vehicle circulation plan. Due to the availability of this information for Pfalznetz, this rail network has been chosen as a study area. Pfalznetz also has a current plan for electrification of the network, which is fitting for comparison with the results of this study. The circulation plan, after sorting out vehicles for which no matching GTFS data could be identified (refer to section 4.5.1), contains 30 vehicles circulating simultaneously (refer to Appendix A, Table 12). They do so on the relevant tracks in the Pfalznetz network (refer to Appendix C, Figure 119). Even though it cannot be visualized at the scale of the map, the data is stored by track so that parallel tracks are treated independently with regards to electrification status and circulating vehicles. The network is comprised of a total 404.5 km of track of which 153.5 km are already electrified.

Due to processing time constraints associated with optimizing the charging infrastructure placement for the entire network, the sensitivity analysis is conducted on a smaller data subset. I chose vehicle circulations 156632707, 156632710 and 156632711 for this purpose. Refer to Table 1, Table 2, and Table 3 for the selected circulation subset. Combined, the selected circulation subset contains trips that both overlap and trips that do not. The vehicle circulations are more complicated in nature than many other vehicle circulations within the circulation plan. While trips do repeat, they do so less than other circulations posing a more complex optimization case. The network relevant to the circulations is also partly electrified thereby also testing the algorithm's ability to account for this fact.

Trip id (GTFS)	**Start station**	**End station**	**Start time**
61978	Pirmasens Hbf.	Landau Hbf.	04:35:00
26742	Landau Hbf.	Pirmasens Hbf.	05:56:00
11079	Pirmasens Hbf.	Kaiserslautern Hbf.	07:32:00
3031	Kaiserslautern Hbf.	Pirmasens Hbf.	08:35:00
47994	Pirmasens Hbf.	Saarbrücken Hbf.	09:33:00
11164	Saarbrücken Hbf.	Pirmasens Hbf.	11:07:00
11079	Pirmasens Hbf.	Kaiserslautern Hbf.	12:41:00
3031	Kaiserslautern Hbf.	Pirmasens Hbf.	13:35:00
54012	Pirmasens Hbf.	Saarbrücken Hbf.	14:33:00
200919	Saarbrücken Hbf.	Zweibrücken Hbf.	16:24:00

Table 1: Vehicle circulation subset used in the sensitivity analysis for vehicle 156632707. The circulation is derived from matching the Pfalznetz circulation plan with a GTFS dataset (refer to section 4.5.1).

Trip id (GTFS)	**Start station**	**End station**	**Start time**
11079	Pirmasens Hbf.	Kaiserslautern Hbf.	06:33:00
3031	Kaiserslautern Hbf.	Pirmasens Hbf.	13:35:00
54012	Pirmasens Hbf.	Saarbrücken Hbf.	14:33:00
11164	Saarbrücken Hbf.	Pirmasens Hbf.	16:07:00
11079	Pirmasens Hbf.	Kaiserslautern Hbf.	17:39:00
3031	Kaiserslautern Hbf.	Pirmasens Hbf.	18:35:00
47994	Pirmasens Hbf.	Saarbrücken Hbf.	19:33:00
200919	Saarbrücken Hbf.	Zweibrücken Hbf.	21:07:00

Table 2: Vehicle circulation subset used in the sensitivity analysis for vehicle 156632710. The circulation is derived from matching the Pfalznetz circulation plan with a GTFS dataset (refer to section 4.5.1).

Trip id (GTFS)	**Start station**	**End station**	**Start time**
179676	Zweibrücken Hbf.	Pirmasens Nord	06:01:00
866084	Pirmasens Nord	Saarbrücken Hbf.	06:38:00
11164	Saarbrücken Hbf.	Pirmasens Hbf.	08:07:00
11079	Pirmasens Hbf.	Kaiserslautern Hbf.	09:41:00
3031	Kaiserslautern Hbf.	Pirmasens Hbf.	10:35:00
47994	Pirmasens Hbf.	Saarbrücken Hbf.	11:33:00
11164	Saarbrücken Hbf.	Pirmasens Hbf.	13:07:00

11079	Pirmasens Hbf.	Kaiserslautern Hbf.	14:41:00
3031	Kaiserslautern Hbf.	Pirmasens Hbf.	15:35:00
47994	Pirmasens Hbf.	Saarbrücken Hbf.	16:33:00
11164	Saarbrücken Hbf.	Pirmasens Hbf.	18:07:00
11079	Pirmasens Hbf.	Kaiserslautern Hbf.	19:41:00
3031	Kaiserslautern Hbf.	Pirmasens Hbf.	20:35:00

Table 3: Vehicle circulation subset used in the sensitivity analysis for vehicle 156632711. The circulation is derived from matching the Pfalznetz circulation plan with a GTFS dataset (refer to section 4.5.1).

The network pertaining to the circulation subset has a combined length of 143.5 km of which 18.2 km are already electrified. The network extends between the four major stations Saarbrücken Hbf., Pirmasens Hbf., Kaiserslautern Hbf., and Landau Hbf. that are all connected to one another through the railway lines intersecting at Pirmasens-Nord (refer to Appendix C, Figure 120). Only the initial part of vehicle circulation 156632707 relates to the tracks connecting to Landau Hbf. Otherwise, the entire circulation subset moves on the tracks connecting Pirmasens Hbf., Saarbrücken Hbf., and Kaiserslautern Hbf.

4.3 Train Vehicle

The spatial distribution of charging infrastructure is optimized according to the operation of a specific train vehicle. The vehicle attributes used in this thesis are not those of any particular vehicle but rather an attempt at an imagined generic BEMU. The relevant specifications of the generic BEMU are determined based on the experience and expertise of the supervisors Sebastian Herwartz and Johannes Pagenkopf. The vehicle specifications are used in the simulations of train energy demand and in the calculations of battery SoE. All vehicle parameters are adjustable but for the purposes of this thesis no sensitivity analysis was performed on the alterations of variables described in Table 4. Some relevant variables of the train vehicle are implicitly included in other variables for the purposes of simplification or when insufficient knowledge is present to presume the exact value of a parameter. This is, for example, the case for the power demand of auxiliaries that is a given as single constant. This is a significant reduction in complexity from real-world calculations of auxiliary power draw. The methodology of this thesis could be conducted on any set of train vehicle parameters.

Parameter	**Abbreviation**	**Value**
Number of seats	None	120
Vehicle length	None	42 m
Static mass	None	97520 kg
Rotating mass	None	4750 kg
Max speed	None	140 km/h
Max deceleration	None	1 m/s2
Max acceleration	None	1 m/s2
Max jerk traction	None	1 m/s3
Max jerk braking	None	1 m/s3
Power efficiency of the gear	eff_gear	0.975
Power efficiency of the motor	eff_motor	0.95
Power efficiency of the battery	eff_battery	0.95
Power efficiency of the AC/DC-converter at the motor	eff_acdc_motor	0.975
Power efficiency of the AC/DC-converter at the catenary	eff_acdc_catenary	0.975
Power efficiency of the DC/DC-converter at the battery	eff_acdc_battery	0.975
Power draw of auxiliaries	aux_demand	60000 W

Max power recouped when braking (by velocity)	max_break_recoup	10000 W
Max power recouped from catenary (absolute)	max_catenary_recoup	900000 W

Table 4: Specifications of the generic battery-electric regional train vehicle used as an example vehicle.

The power flow in the vehicle mechanics between the catenary, the battery, the motor, and the auxiliaries is important for calculating the change in battery SoE. Figure 3 visualizes how relevant variables in Table 5 relate conceptually for the purposes of power and energy calculations.

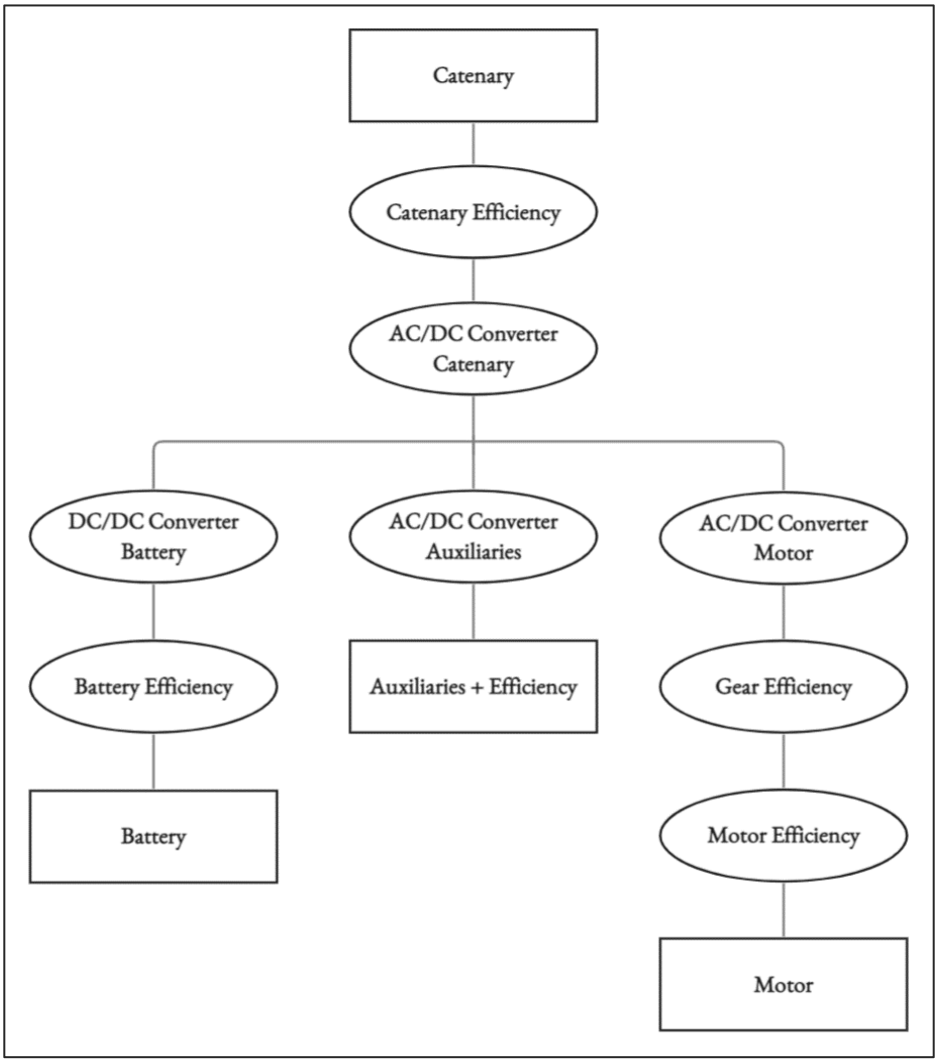

Figure 3: Conceptual overview of the relationship between elements in the train vehicle relating to the calculation of power and energy flows.

Based on the power at wheel (the power required, or potentially recouped during braking, to operate the vehicle at the point of the wheel, referred to as Pwhl), the potential change in the battery SoE (referred to as ebatt) at the specific point is calculated using the formulas and relationship provided below. All calculations of the potential change in the battery SoE uses the Pwhl and p_acdc_motor. p_acdc_motor refers to the power at the motor based on the Pwhl after accounting for the relevant efficiencies. The selection of the functions depends on the presence of catenary and the mode of operation of the motor. Catenary can either be present or not. The motor can either be driving (actively accelerating), stationary/coasting (no motor activity), or braking (actively decelerating). This results in the matrix presented in Table 5. The vehicle requires power to accelerate and can charge through catenary or recouperation from braking.

Variable	Catenary	Mode of Operation
ebatt_driving_catenary	Present	Driving
ebatt_stationary_catenary	Present	Stationary/Coasting
ebatt_braking_catenary	Present	Braking
ebatt_driving	Not present	Driving
ebatt_stationary	Not present	Stationary/Coasting
ebatt_braking	Not present	Braking

Table 5: Matrix describing the names of ebatt calculation formulas to be used depending on the presence of catenary and the motor mode of operation.

p_acadc_motor is calculated from the Pwhl in two separate directions depending on whether the vehicle is driving or not. If the vehicle is driving, the power must move flow towards the motor and therefore additional power is required as a factor of the relevant efficiencies to account for the power loss. If the vehicle is not driving, then power flows towards the battery and is reduced by the relevant efficiencies. In this case, the maximum amount of power that can be recouped is limited by the max_break_recoup multiplied by the vehicle velocity.

$if\ Pwhl > 0$:

$$p\_acdc\_motor = \frac{Pwhl}{eff\_gear * eff\_motor * eff\_acdc\_motor}$$

$Otherwise\ if\ Pwhl \leq 0$:

$$p\_acdc\_motor = \max_{(max\ \_break\_recoup * -1) * velocity} Pwhl * \ eff\_gear * eff\_motor * eff\_acdc\_motor$$

If the vehicle is not under catenary and driving, the ebatt is the power draw of the auxiliaries and the motor increased by the relevant efficiencies. At the battery, this value must be negative. It is converted from W to energy in kWh.

$$ebatt\_driving = \frac{\frac{\frac{p\_acdc\_motor + aux\_demand}{eff\_dcdc\_battery * eff\_battery * -1}}{1000}}{3600}$$

If the vehicle is not under catenary and stationary, the ebatt is the power draw from the auxiliaries increased by the relevant efficiencies. At the battery, this value must be negative. It is converted from W to energy in kWh.

$$ebatt\_stationary = \frac{\frac{\frac{aux\_demand}{eff\_dcdc\_battery * eff\_battery * -1}}{1000}}{3600}$$

If the vehicle is not under catenary and braking, then the calculation depends on whether the total of the recouped power at the motor and the power draw of the auxiliaries is positive or negative. If it is negative, then power flows to the battery and is reduced by efficiencies, if it is positive, then power flows from the battery and must be increased by the efficiencies. In the latter case, it must also be converted to a negative value as it is a loss from the battery's perspective. The result is converted from W to energy in kWh.

$if\ p\_acdc\_motor + aux\_demand < 0$:

$$ebatt\_braking =$$

$$\frac{\frac{(p\_acdc\_motor + aux\_demand) * eff\_dcdc\_battery\ * eff\_battery}{1000}}{3600}$$

$Otherwise\ if\ p\_acdc\_motor + aux\_demand \geq 0$:

$$ebatt\_braking = \frac{\frac{\frac{p\_acdc\_motor + aux\_demand}{eff\_dcdc\_battery\ * eff\_battery * -1}}{1000}}{3600}$$

If the vehicle is under catenary and driving, then the direction of the power flow is determined by whether the power gained from the catenary is higher than draw from the motor and auxiliaries. If this value is negative, then power flows flow the battery and must be negative and increased by efficiencies. If the value is positive, then power flows to the battery and must be reduced by relevant efficiencies. The result is converted from W to energy in kWh.

$if\ max\_catenary\_recoup\ - aux\_demand\ -\ p\_acdc\_motor < 0$:

$$ebatt\_driving\_catenary = \frac{\frac{\frac{(max\_catenary\_recoup * eff\_acdc\_catenary) - aux\_demand - p\_acdc\_motor}{eff\_dcdc\_battery * eff\_battery * -1}}{1000}}{3600}$$

$Otherwise\ if\ max\_catenary\_recoup\ - aux\_demand\ -\ p\_acdc\_motor \geq 0$:

$$ebatt\_driving\_catenary = \frac{\frac{((max\_catenary\_recoup * eff\_acdc\_catenary) - aux\_demand - p\_acdc\_motor) * eff\_dcdc\_battery * eff\_battery}{1000}}{3600}$$

If the vehicle is stationary under catenary, then the power flows to the battery. This is equal to the power gain from the catenary minus the power draw from the auxiliaries reduced by relevant efficiencies. This is converted from W to energy in kWh.

$$ebatt\_stationary\_catenary = \frac{\frac{(max\_catenary\_recoup\ * eff\_acdc\_catenary - aux\_demand) * eff\_dcdc\_battery\ * eff\_battery}{1000}}{3600}$$

If the vehicle is braking under catenary, then power flows to the battery from both the motor and the catenary. p_acdc_motor is inverted to positive as this should be an increase from the perspective of the battery. These are subject to the relevant efficiencies and are subtracted the demand from auxiliaries. The result is converted from W to energy in kWh.

$$ebatt\_braking\_catenary = \frac{\frac{((max\_catenary\_recoup\ * eff\_acdc\_catenary) + (p\_acdc\_motor * -1) - aux\_demand) * eff\_dcdc\_battery * eff\_battery}{1000}}{3600}$$

4.4 Data

The methodology builds on six distinct sets of data: a vehicle circulation plan, a GTFS dataset, OSM data, route geometries, a DTM, and simulated vehicle power demand at wheel. Each dataset constitutes a necessary input for the optimization of charging infrastructure. This section describes each dataset and its core function in the methodology.

4.4.1 Vehicle Circulation Plan

Pfalznetz is chosen as the study area specifically because a vehicle circulation plan was available for this particular subnetwork. The vehicle circulation plan provides information on the order of which each vehicle initiates and completes trips through a day. This information is needed to calculate the battery SoE throughout an entire day. This is both to account for the variable energy requirement between trips and to account for the amount of time (and position in space) the vehicle spends between trips. Some vehicles mostly go back and forth on the same route whilst others have more diverse circulations.

The Pfalznetz circulation plan contains a unique id for each vehicle and a unique id for each trip. These ids are internal to the circulation plan and do not match the id within any other dataset. The formatted and linked circulation plan is appended in appendix A, Table 12 (refer to section 4.5.1). It has been manually transcribed as a comma-separated file to extract the data for further data processing. Information on each trip, specifically when it begins and which stops it passes through, is extracted from a separate file.

4.4.2 General Transit Feed Specification (GTFS)

The DLR has purchased a GTFS dataset for the entirety of Germany. As does any dataset in the GTFS file format [27], this contains information on the public transit trips of Germany. Each unique trip in time and space has an identification number (Trip_id) that is associated with the stops (stored in stops.txt) through which it passes at given times (stored in stop_times.txt). The dataset is used to gather information on the movement of the vehicles during their circulation. An example data extract is provided in Table 6 and Table 7.

Stop_name	Stop_id
Dabendorf	10013
Salzburg Hbf	10014
Asselheim	10015
Wörth(Main)	10016

Table 6: An example of stops.txt from the GTFS dataset.

Trip_id	Arrival_time	Departure_time	Stop_id	Stop_sequence
12187	15:24:00	15:24:00	11599	0
12187	15:25:00	15:25:00	13050	1
12187	15:26:00	15:26:00	5895	2
12187	15:27:00	15:27:00	4811	3

Table 7: An example of stop_times.txt from the GTFS dataset.

4.4.3 Routed Geometries

For each spatially identical set of trips in the Pfalznetz vehicle circulation plan, I manually created line geometries in the appropriate direction matching their path through space. The routing of these geometries follow the railway network using the BRouter service [41]. The correct tracks for each trip for sections with multiple tracks is assumed based on presumption that train vehicles always keep right. Each geometry created this way is assigned a unique id.

4.4.4 OpenStreetMap

OSM data is used to enrich the routed geometries with information on electrification and max speeds. The max speeds are used in the simulations of vehicle power at wheel while the current state of electrification is used in its unaltered state throughout the genetic algorithm optimization. The routed geometries are enriched with OSM data through a modified script originally developed by J. Polster [42]. OSM data on the locations of stations is also used in the network segmentation process (refer to section 4.5.2). OSM data is downloaded for free from geofabric [43].

4.4.5 Digital Terrain Model

This thesis uses a 30x30 m resolution DTM extracted from NASA open data [44]. The DTM data is used only as input for the DLR developed MatLab tool simulating vehicle power at wheel during operation. The elevation is extracted along each routed geometry representing the path in space traversed for each spatially identical set of trips.

4.4.6 Simulations of Vehicle Power at Wheel

Simulations of the power at wheel of the train vehicle at second intervals are used in the genetic algorithm optimization in the calculation of the potential change in battery SoE at each second of its circulation. The simulations are run by the DLR using an internally developed tool built in MatLab. Conceptualizations of the tool are presented by Schenker, M. et al. [45], [46] but the version and model results employed in this thesis are not a direct result of these publications in particular.

The tool uses information on elevation gradients, track speed limits, and specifications on the train vehicle to simulate the power at wheel required or possibly recouped at each second of the timetable of the trip to satisfy the trip timetable using the least power. The tool optimizes the driving strategy of the vehicle so that the vehicle can satisfy the timetable in the most power efficient manner. The information on elevation gradients and track speed limits are extracted and fed to the tool from a DTM and OSM respectively. Timetable information is derived from the GTFS dataset. The data extraction and formatting process is based on amended scripts originally developed by J. Polster [42]. A simulation is run for each spatially identical set of trips in the GTFS dataset. A set of trips is spatially identical when they pass through the same stations (but may still vary in time).

Different train vehicle driving strategies impact the way in which the optimal composition of modes of operation is determined by the tool. For this thesis, the tool is configured to include coasting (allowing the wheels to roll without motor interference). This is in addition to braking (active deceleration via the motor), constant speed (maintaining a given target speed with motor assistance) and accelerating (actively accelerating with the motor).

4.5 Data Processing and the Genetic Algorithm Structure

The section describes all steps transitioning from initial input data to an optimized spatial distribution of charging infrastructure for any given scenario. Some processing steps are only conducted once and are used as input for all model runs whilst others are performed at the initiation or during each model run. This difference is explicitly stated throughout the section. Most of the method is written and executed in Python but few preprocessing steps are performed with QGIS and/or completed with manual edits.

4.5.1 Connecting the Vehicle Circulation, the GTFS, and the Routed Geometries

The vehicle circulation plan, the GTFS dataset and the routed geometries are manually connected to one another in a final circulation table. This is done a single time for the entire dataset. An example of this data format is provided in Table 8.

The transcribed Pfalznetz trip id is connected to the GTFS dataset candidate id through manual search of matching trips. The candidate ids are shared ids between trips in the GTFS dataset that share stops. The candidate ids were identified by J. Polster [42]. A routed geometry is manually created (see section 4.4.3) for each spatially unique path through space in the circulation plan and associated with the correct candidate id from the GTFS dataset.

Vehicle number	Trip id (Pfalznetz)	Trip id (GTFS)	Geometry id	Sequence	Start time
156632710	64034	11079	5	0	06:33:00
156632710	64013	3031	6	1	13:35:00
156632710	65011	54012	13	2	14:33:00
156632710	65014	11164	11	3	16:07:00
156632710	64020	11079	5	4	17:39:00

Table 8: An exemplary extract of the pre-processed vehicle circulation plan linking to the GTFS dataset and the routed geometries.

Not all parts of the Pfalznetz vehicle circulation could be successfully matched with the GTFS dataset. Some trips in the Pfalznetz circulation did not exist in their entirety but had a close match (for example one smaller stop missing). If a close match could be identified, this is used as a substitute. The substitution process, even though it reduces the accuracy of the examined circulation plan, has not had an effect on the feasibility of the vehicles' ability to complete the prior GTFS trip before initiating the next GTFS trip at the time given in the circulation plan. Other trips which did not exist and did not have a close match in the GTFS data have been omitted. If these were at the end or beginning of a vehicle circulation, they have been removed from the circulation. If they were present in the middle of a vehicle circulation, the entire vehicle circulation has been omitted. This was only the case for a few vehicle circulations. The edited and matched Pfalznetz vehicle circulation is doable within the time plan given by the matched GTFS dataset.

4.5.2 Network Segmentation

The entire network is simplified and segmented before initiating models run. At the start of an algorithm run, the relevant parts of the segmented network are selected so that only segments that relate to all trips conducted in the selected circulation are included in the algorithm run. The selection process is programmed as part of the algorithm pipeline.

Initiating the segmentation, all routed geometries are simplified into a single network with no overlapping parts. Each segment includes information on which trips pass through it. The segments split when they differ in which trips pass through. This first part of the segmentation is performed manually in QGIS due to spatial data inconsistencies making programing a standardized approach unfeasible. Secondly, all segments are split by current electrification status (derived from OSM). No segment can be both electrified and non-electrified. Thirdly, the segments are split by length. This process is two-part. To ensure that station segments center on the station, the segments are first split by station buffers created from the OSM data. The radius of the buffer is set to 250 m, creating segments centered on each station with a length of 500 m. Since trips vary on their exact end point in space, it is important to prevent segments from splitting in the middle of a station. This ensures that different trips are registered as stopping on the same station segments. As a last step, the remaining non-electrified segments are split by max length so that no segment is longer than 500 m. A smaller segment length increases the granularity with which the algorithm can place electrification but does so at the cost of processing time. The 500 m chosen for this thesis represents a compromise between the two. Already electrified segments do not need to be split as they are not evaluated in the genetic algorithm. This reduces processing time. An illustrative image of the resulting segmented network is provided on Figure 4 while a conceptual illustration of the splitting process is provided on Figure 5.

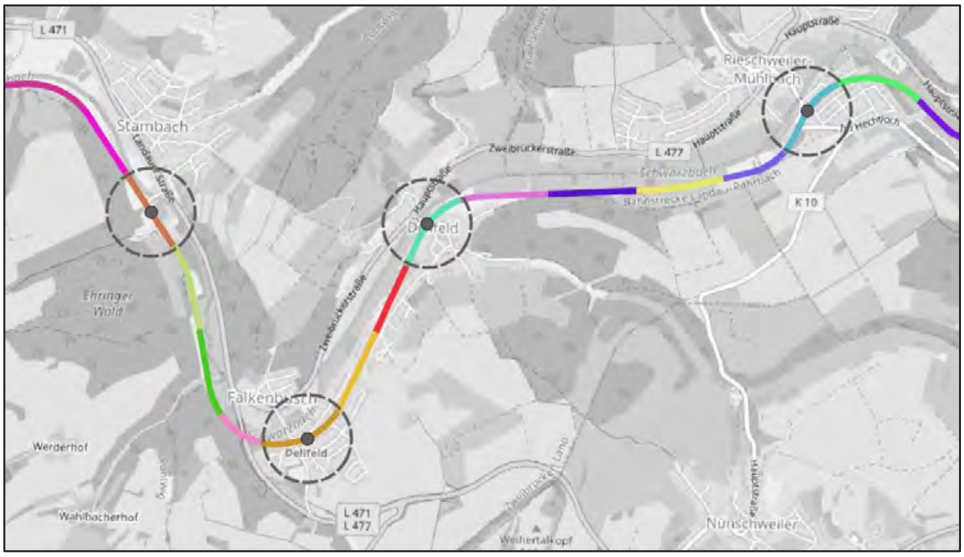

Figure 4: A cropped image of the segmented network in QGIS that is split by relevant attributes, station buffers and length.

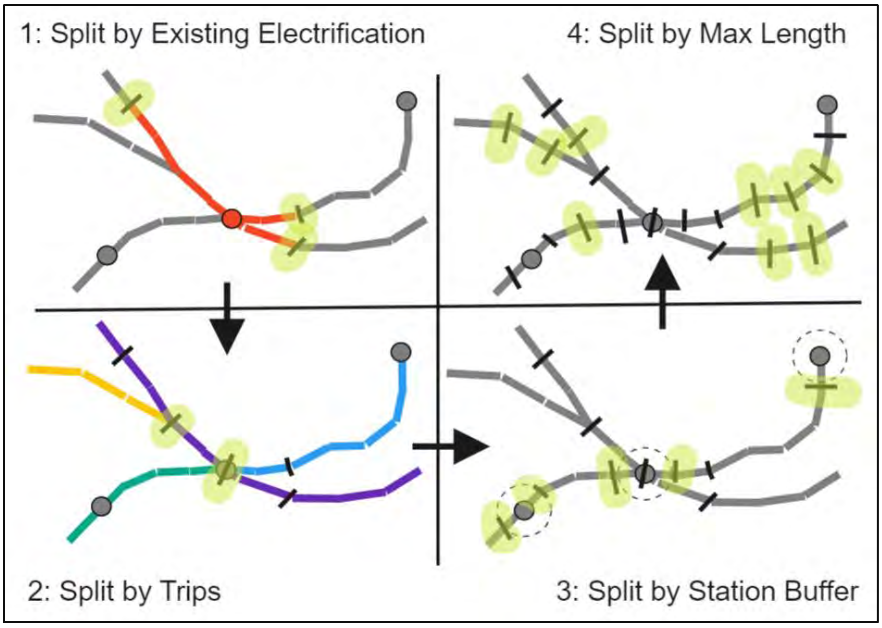

Figure 5: A conceptual illustration of the network segmentation process including its four main steps. The network is split further at each stop of the process.

The sequence through which a given trip passes through the segments is identified after the network has been segmented. I create points at 1 m intervals along each routed geometry and evaluate for each point which buffered segment captures it. The order in which points are captured by different segments is equal to the sequence.

4.5.3 Calculations of Potential Change in Battery State of Energy

Before the initiation of a model run within the algorithm pipeline, relevant simulations of vehicle power at wheel are formatted and appended as a single simulation for each vehicle in the selected circulation. For each spatially unique set of trips in the selected circulation, the relevant information from the simulation of vehicle power at wheel is extracted and the potential change in battery SoE for both catenary and without catenary is calculated for each simulated second using the formulas described in section 4.3. Whether catenary is present or not is determined within the genetic algorithm. The simulations are then summed by segment according to the sequence in which the trip passes through them so that each row represents the time spent and the potential change in battery SoE at that segment. The calculated potential changes in battery SoE for each second of operation are allocated their respective network segment using the distance travelled by the vehicle in the simulation combined with the sequence in which the vehicle passes through each segment. Refer to Figure 6 for a conceptual illustration of these calculations.

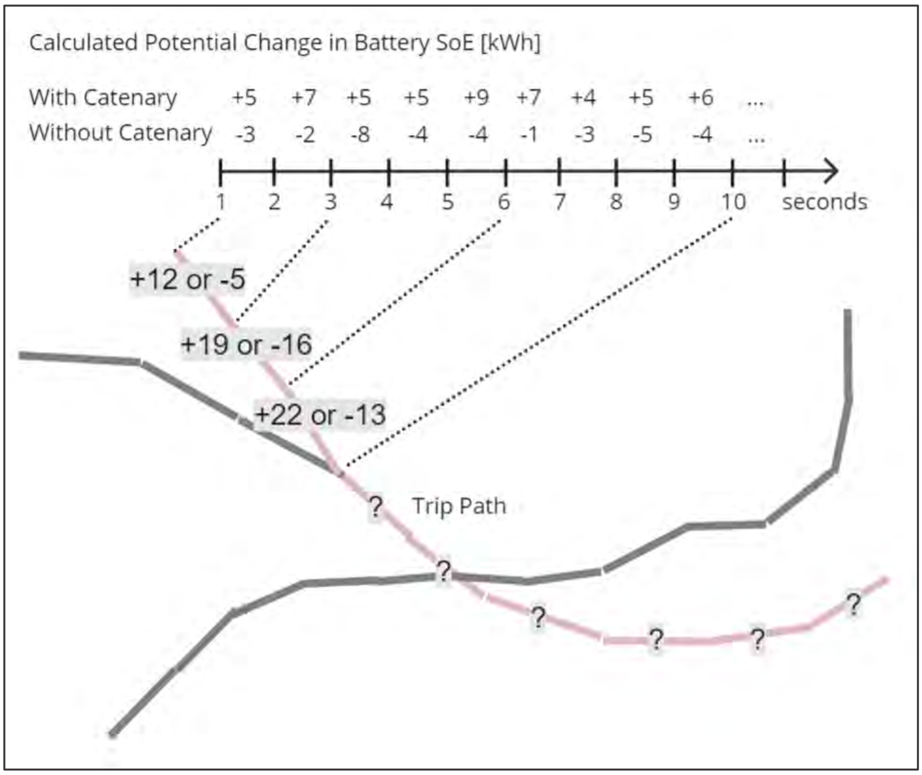

Figure 6: A conceptual illustration of how potential change in battery SoE in seconds is summed by segment as represented by two mutually exclusive changes; with or without catenary. The values in the figure are not representative of actual values.

All formatted simulations pertaining to a single vehicle are appended together in order to produce a single energy simulation for that vehicle's circulation. This also includes calculating the time spent between trips and the potential change in battery SoE resulting from this time. In many cases, the amount of time spent between trips can be the most efficient time/place to charge the vehicles. An example of the final simulation format is provided in Table 9.

Segment id	**Change in battery SoE, without catenary [kWh]**	**Change in battery SoE, with catenary [kWh]**	**Time on segment [s]**	**GTFS id**	**timestamp**
46	-0.094753	1.499122	7	61978	04:35:07
47	-4.560344	4.881358	30	61978	04:35:39
499	-0.177505	5.821841	25	61978	04:36:02
500	-2.056685	5.235622	25	61978	04:36:27

Table 9: An example of the final simulations input format to the genetic algorithm. Potential change in the battery SoE is calculated for each segment in all trips of the vehicle circulation. This table is calculated for each vehicle. Some columns used in the calculations have been excluded for clarity.

4.5.4 Genetic Algorithm Optimization

The genetic algorithm takes the selected network and the simulated vehicle circulation – both of which are formatted within the algorithm pipeline before initiating the algorithm. Each run takes the parameter inputs shown in Table 10 (in addition to the vehicle parameters in Table 4). Most of these remain constant across the model runs of the thesis but the script is flexibly designed such that they may be altered to accommodate different vehicles, costs, etc. The script for the genetic algorithm is a modified version of that provided by the DEAP Python package [47]. The package structure forms the framework for the implementation of the genetic algorithm.

Parameter Name	**Value**	**Brief Explanation**
Vehicle numbers	N/A	A list of vehicle numbers from the circulation to optimize based on.
Max SoE [kWh]	N/A	The maximum energy capacity of the battery.
Min SoE [kWh]	50	The threshold for which the battery is considered too low in energy and the penalty is applied.
Substation Cost [eur/unit]	N/A	The estimated cost assigned to each substation serving a distinct catenary island.
Catenary Cost [eur/m]	610	The cost of catenary.
Low SoE Penalty [eur/m]	12600	The cost assigned for each meter any vehicle in the selected circulation is below the Min SoE [kWh] threshold.
Substation Connection Distance [m]	25	The distance with which segment can share substations (per segment, therefore 50 m between segments).
Population Size	200	The size of the populations of each generation.
Selection Function	Tournament	The function used for the selection process.
Tournament Size	5	The tournament size (if the tournament function is selected).
Mutation Function	Flip Bit	The function used for the mutation process.
Average Gene Mutation Amount	3	The average amount of segments for which the electrification status is flipped during mutation.
Mutation Probability	0.5	The probability of each individual in a population to be mutated.
Crossover Function	Two-Point	The function used for the crossover process.
Crossover Probability	0.5	The probability of each individual in a populated to take part in a crossover.
Max Runtime [s]	N/A	The maximum runtime of the algorithm after which it terminates.
Max Generations	N/A	The maximum number of generations of the algorithm after which it terminates.

Table 10: Input parameters for the genetic algorithm. Values marked with N/A are variable between the model runs conducted within this thesis.

At the initiation of the genetic algorithm, the individuals are encoded as lists of binary values with a length equal to the amount of non-electrified segments within the selected network for that model run. These are the segments for which the genetic algorithm should determine the optimized configuration of electrification. Each gene within the individual represents the state of electrification for a specific currently non-electrified segment (refer to Figure 7). During the model development, I experimented with population seeding (starting with a non-random composition of initial individuals) based on average time spent on segments. Generally, segments on which vehicles spend larger amounts of time provide more cost-efficient opportunities for charging. As this did not lead to improved processing times or results, population seeding is abandoned in favor of randomizing the genes of the individuals in the initial population.

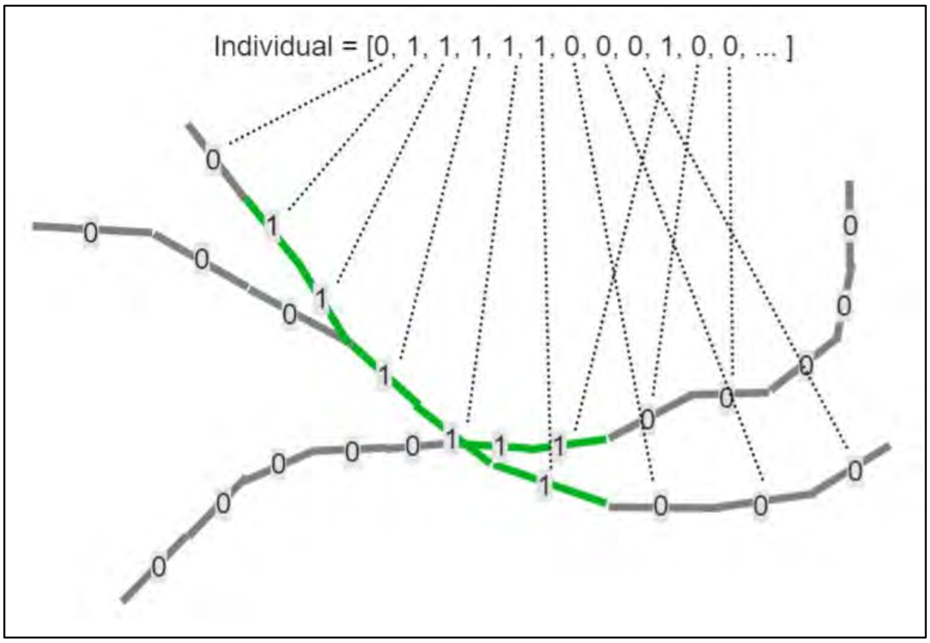

Figure 7: A conceptual illustration of the binary list encoding of individuals within the genetic algorithm. The genes in an individual each represent the state of electrification at a network segment. The order of the genes does not necessarily represent the order of segments in space.

Each individual in each population is evaluated using the fitness function. The fitness is based on the sum of two values: the infrastructure costs and the penalty for insufficient battery SoE at any part of the vehicle circulation. The infrastructure cost is estimated using the input infrastructure cost parameters combined with the total length of electrified track and the number of substations. The number of substations is identified by dissolving a buffer of electrified segments and counting the number of resulting polygons. This number will be equal to the amount of electrification islands each serviced by a substation. A conceptual illustration of catenary islands and the counting of substations is visualized on Figure 8. The size of the buffer is equal to the input parameter substation connection distance controlling the range a substation can service segments that are not directly connected. This is set to 50 m (2x25 m radius) which I evaluate to be the minimum value necessary to allow for the sharing of substations between all tracks on the largest station in the study area (Kaiserslautern Hbf.).

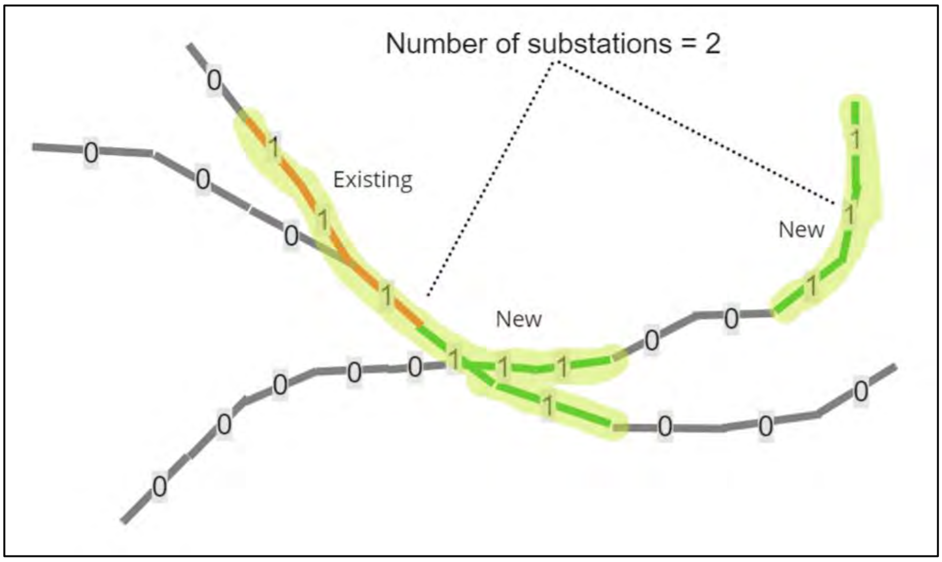

Figure 8: A conceptual illustration of catenary islands and how they relate to the number of substations. Catenary islands are isolated contiguous stretches of electrification each requiring its own substation. Substations already exist at existing electrification.

The estimated infrastructure costs are added to the penalty assigned for insufficient battery SoE. According to the electrification composition encoded within the evaluated individual, the potential change in battery SoE for each traversed segment in the simulated vehicle circulations is selected (either under or not under catenary). The battery SoE for each vehicle in the circulation throughout their course is then calculated from the selected potential changes accounting for the maximum battery capacity that cannot be exceeded. For each meter in any simulated vehicle circulation the battery SoE is under the threshold defined by the minimum battery SoE parameter, the penalty is added to the fitness value. The penalty is set to 20 times the cost of catenary per meter, ensuring that it is always cheaper to install catenary than to leave insufficient charging infrastructure.

The fitness evaluation step is, by far, the largest processing power draw of the algorithm. I implemented multicore processing at this step of the algorithm allowing all CPU cores operate in parallel calculating the fitness for smaller batches of the population. For the hardware used in this thesis, this sped up processing time almost directly proportional to the amount of CPU cores.

After all individuals in the population have been evaluated, they are selected, crossed-over and mutated using the methods for this defined in the genetic algorithm parameters. For the model runs of this thesis, I chose the selection of individuals to pass on to the next generation to be identified by the tournament method with a tournament size of 5. The algorithm selects the best individual (lowest fitness value) from random batches of 5. It continues this process until the population size is selected. This allows for the same individual to be selected repeatedly if it continues to win tournaments where it randomly participates. This selection method is both stochastic limiting the algorithm's tendency to gravitate to local minima while almost always allowing the best individuals of a generation to pass on to the next generation.

Half of these individuals are then crossed over (at random) producing new individuals that are mixes of their parents. I opted for one of the simpler crossover methods within genetic algorithm research: the two-point crossover. This picks two points at random in the individual's' genes at which to cut and splice them with those of the other parent to produce their children/offspring (refer to Figure 9). This simple function is picked in order to limit the mutating impact of the crossover method given the nature of this particular problem. The segments represented in the genes are not in spatial chronological order (no such thing exists in a complicated network) but do tend to be near or next to each other in space (refer to Figure 7). Consequently, when crossover is performed, stretches of catenary can be accidentally split, resulting in additional substation costs and thus poorly performing individuals. Two-point crossover limits this unintended consequence of crossover by only allowing two points to be chosen. During the parameter tuning process, I experimented with single-point crossover to further simplify the crossover process. This reduced the efficiency by which the algorithm reached optimal outcomes and is therefore discarded in favor of the two-point crossover. This serves as a compromise between efficient stochasticity and limited amounts of poorly performing offspring.

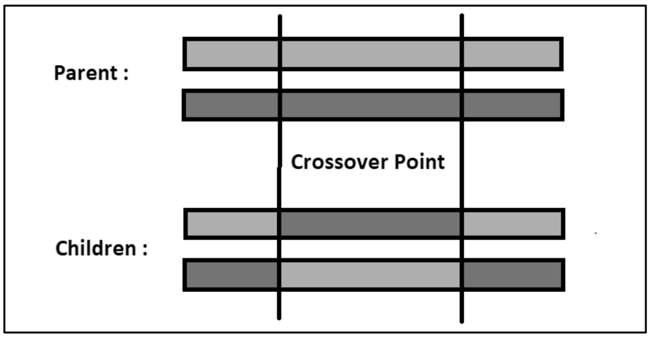

Figure 9: A conceptual illustration of the two-point crossover applied in the genetic algorithm [48].

Following the crossover, half of all individuals undergo mutation. This mutation is set to flip bit that is standard for individuals with binary encoding. Flip bit randomly flips the value of each gene in the individual according to the gene mutation probability (refer to Figure 10). Here, the methodology of this thesis deviates slightly from how mutation chance is often handled in the field of genetic algorithms. Firstly, the probability of each individual undergoing mutation is relatively high (50%). For this problem, the main way in which the algorithm explores the problem space is through mutation. The crossovers optimize individuals in the existing population but do not explore well the potential locations to add or remove electrifications. This exploration is performed by the mutation, and it is therefore important for mutations to occur often. Lowering the mutation chance would still allow for the process to occur but it would do so significantly slower. Due to time processing constraints, this is not viable.

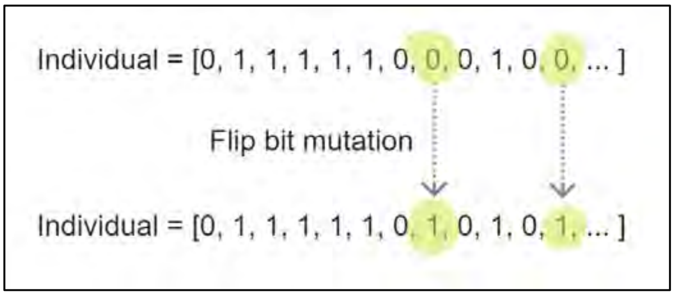

Figure 10: A conceptual illustration of flip bit mutation. All genes in the mutated individual have a probability to flip bit.

The high mutation probability is combined with a low probability for each gene in the individual to flip bit. The probability for each gene to flip bit when an individual undergoes mutation is normally set as a fixed value, however, in the construction of this genetic algorithm pipeline, the probability for each gene to undergo mutation is instead set relative to the size of the individual. The gene mutation probability is determined so that a given number of segments are, on average, flipped during mutation irregardless of the size of the selected network. With an average number of flipped segments set to 3, the value chosen for the model runs conducted in this thesis, the probability of a gene flipping during mutation in an individual with a length of 600 would be 0.5% while it would be 1% for an individual with a length of 300. Across series of populations, the number of flipped bits within the mutated individuals distribute according to a normal distribution visualized on Figure 11. It is important to keep the precision of the gene mutation small across model runs – even those conducted on large data subsets. For this particular problem, if the gene mutation probability is too high, then the algorithm fails to identify better problem solutions once it reaches a half-optimized state. When too many bits flip at once, the likelihood that at least one of them flips from 0 to 1 in a space without electrification is high. This would result in a catenary island with a high cost and thus a poorly performing individual. Setting a consistently low gene mutation probability allows the algorithm to explore the problem space and find better solutions even when a good solution is already identified.

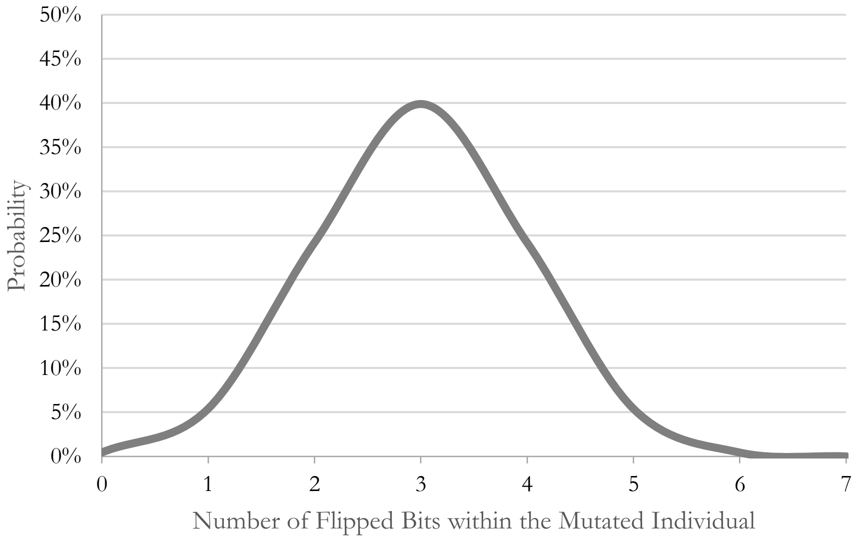

Figure 11: Probability density curve for the number of flipped bits within mutated individuals regardless of individual length with a mean number of flipped bits set to 3.

After mutation, the process begins anew iterating on the new population created. The termination criteria are either set by a generation limit or by processing time depending on the model run (refer to section 5.3). When the model terminates, it outputs the identified optimal configurations of charging infrastructure as a geopackage, and the model run parameters, results, and simulated energy calculations in an Excel-file.

4.5.5 Post-Processing

All model runs are subject to post-processing to rectify a specific type of predictable undesirable outcome that is a direct consequence of the structure of the genetic algorithm approach. When a segment has a length less than the distance a substation can be shared between segments (50 m for the model runs conducted in this thesis) and is surrounded by electrified segments on either side, the algorithm will often not electrify the given segment even though it leads to a non-contiguous stretch of electrification. Given how the problem is defined, contiguity is maintained since the area is still treated as a single catenary island. To solve this specific issue and output more reasonable spatial distributions of catenary, all segments satisfying these specific criteria are set to electrified post-process and the vehicles' battery SoE are re-calculated accordingly. This only occurs for a few segments and does so for segments with neglectable impact on the overall SoE.

4.6 Model Runs and Sensitivity Analysis

In order to understand the workings and performance of the thesis methodology, I perform a limited sensitivity analysis on key input parameters (presented in Table 11). To limit processing time, the sensitivity analysis is performed on the reduced study area described in section 4.2. Even with this reduced dataset, a single model run still requires approximately an hour of computation on available hardware (to reach 1000 generations by which time evolution has mostly stagnated).

To examine the impact of infrastructure costs on the optimal spatial distribution of charging infrastructure, I adjust the cost of substations relative to that of catenary. This either incentivizes or disincentivizes the quantity of catenary islands relative to extending existing catenary. I also examine the battery size to investigate how energy needs of the BEMU impact the placement of charging infrastructure. Each model run is conducted thrice to examine the extent of stochasticity in the final results and tendency of the genetic algorithm to approach local minima.

In addition to the model runs for the purposes of sensitivity analysis, I also perform a model run for the entire study area using the standard input parameters. The single model run for the entire study area is not limited by number of generations but instead by computation runtime. It is not needed to compare the result to that of other model runs as only one model run is feasible given the computation time constraints of the thesis. Therefore, I do not constrain the optimization to a given number of generations. Instead, the optimization is terminated after 24 hours.

Model Run Name	**Study Area**	**Battery Capacity**	**Substation Cost**	**Number of Runs**	**Approx. Processing Time**
Entire Pfalznetz	Complete	400 kWh	8000000 €	1	24 h
Small battery	Reduced	300 kWh	8000000 €	3	3 h
Standard battery	Reduced	400 kWh	8000000 €	3	3 h
Large battery	Reduced	500 kWh	8000000 €	3	3 h
Cheap substation	Reduced	400 kWh	4000000 €	3	3 h
Standard substation	Reduced	400 kWh	8000000 €	3	3 h
Expensive substation	Reduced	400 kWh	12000000 €	3	3 h
Total	**N/A**	**N/A**	**N/A**	**19**	**42 h**

Table 11: All model runs conducted within the thesis excluding tuning of the genetic algorithm parameters. The table presents only the variables that change between model runs. All other input variables remain constant.

5. Results

This section presents, highlights, and explores key results relating to the model runs conducted within the scope of the thesis. Before exploring the model outputs, I shed light on the algorithm processing itself and the converge towards optimized distributions of charging infrastructure. Hereafter, the section presents the key results relating to the model run "entire Pfalznetz" serving as the core results of the thesis. The model runs relating to the sensitivity analysis are finally explored to increase the understanding of the model functions. All results are visualized in Appendix B, to which this section continuously refers. Appendix B also contains a large array of results figures that are not directly mentioned within the thesis. Appendix C contains additional map visualizations also referred to in this section.

5.1 Algorithm Processing and Convergence

All conducted model runs, both those within the sensitivity analysis and for those conducted for the entire Pfalznetz circulation plan, follow a similar pattern of convergence towards an optimal solution during the model run. For the complete model run, the final solution is identified at generation 2328. The remaining 172 generations that transpired during the 24-hour model run did not successfully identify a more optimal solution than that already identified (refer to Figure 12). It may be that the model could have identified an even better solution had it been allowed to run for longer. This also applies to all sensitivity analysis model runs. However, given the fact that only negligible progress was made from generation 651, it is unlikely that a longer model runtime duration would have resulted in a solution that would be improved in any significant way.

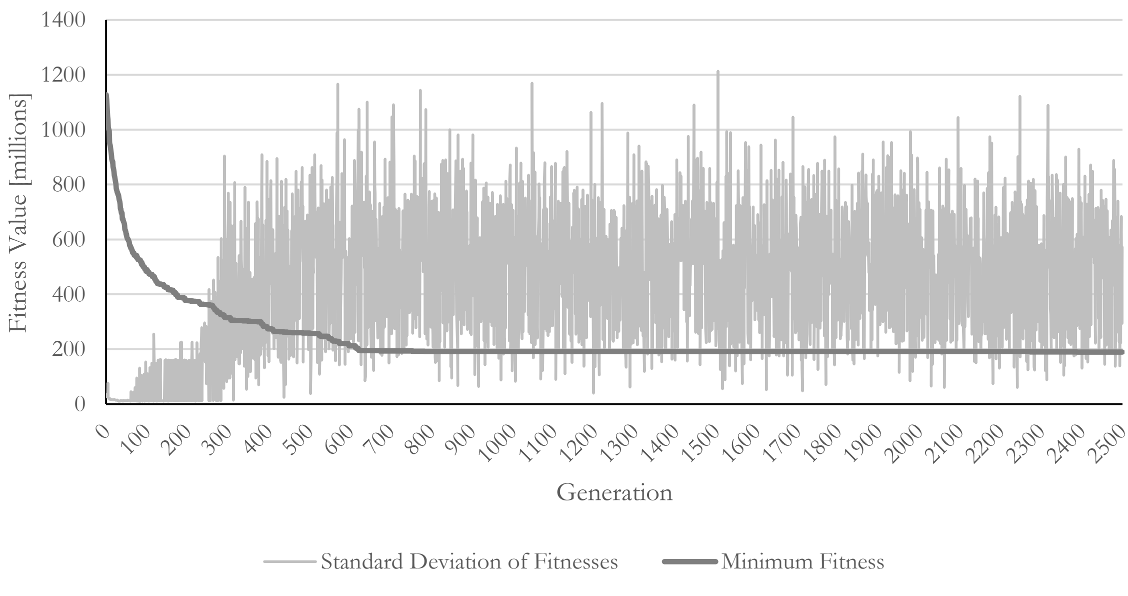

Figure 12: The minimum fitness (the fitness of the best performing individual) and the standard deviation of fitnesses within each population for model run "entire Pfalznetz".

For the sensitivity analysis, there is some variation in how swiftly the algorithm converges on the final output. Most model runs converge near generation 400 and no model run sees further improvement in the fitness of the most fit individual after generation 650. Figure 13 for model run "standard battery" serves as an example of how the model runs on the reduced data subset generally converge on an optimal solution. This is due to the natural stochastic variation in how the genetic algorithm functions. No model run indicates that a notably better solution could have been identified with longer processing time.

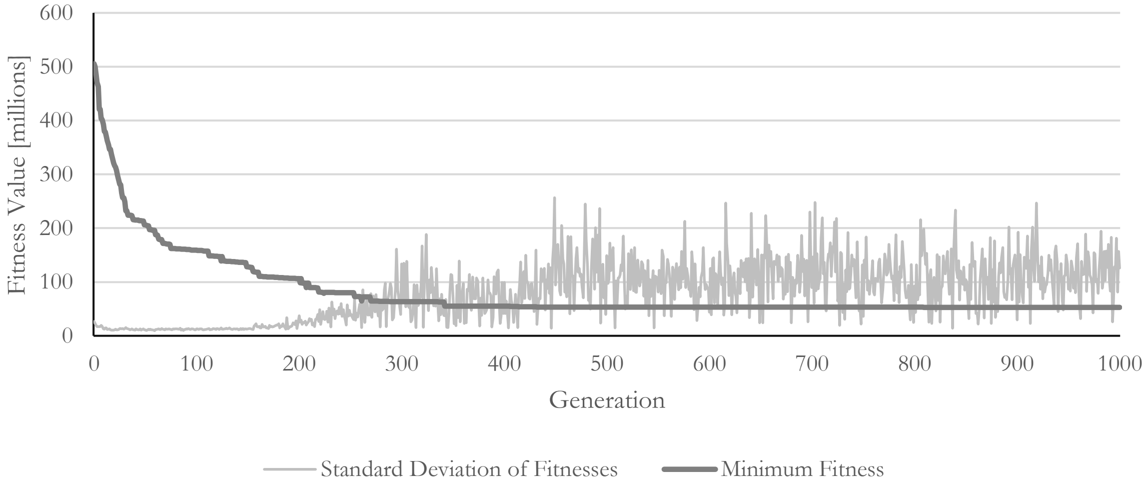

Figure 13: The minimum fitness (the fitness of the best performing individual) and the standard deviation of fitnesses within each population for model run "standard battery 1".

For all model runs, the standard deviation of the fitnesses of the population individuals rises as the best identified solution approaches the optima (regardless of whether this is local or global). This occurs when the vast majority of mutated and crossed individuals in a new population perform considerably worse than that of the best performing individual. When the amount and distribution of new electrification is limited to such an extent that it is nearly unviable to remove any of it without incurring the penalty for insufficient SoE for one or several vehicles, nearly all subsequently created individuals through mutation/crossover will perform poorly. They will either be penalized for insufficient SoE or, if they instead added electrification, have a considerably higher infrastructure cost most likely associated with the construction of new substation infrastructure.

5.2 New Charging Infrastructure in Pfalznetz

The optimized distribution of electrification infrastructure proposed by the model run "entire Pfalznetz" proposes a total 12.9 km additional catenary serviced by an additional 8 substations. The total cost of this additional infrastructure estimated to be €72.4 million, of which substation costs make up €64 million and catenary €8.4 million. As such, the majority of the cost for new the electrification, as constructed within this thesis, is mostly derived from the substations and not from the catenary. New substations are necessary to allow for the construction of catenary in areas that are distinct from existing electrification but still require electrification to be constructed to allow BEMU operation.

The model run proposes new islands of electrification near Kusel, between Kaiserslautern and Lauterecken-Grumbach, at Zweibrücken Hbf., at (and near) Pirmasens-Nord, between Pirmasens-Nord and Landau, and west of Bad Bergzaberg (refer to Appendix B, Figure 17). With this proposed additional electrification supplementing the existing electrification no vehicle dips below the minimum battery SoE threshold (50 kWh) throughout their circulation (refer to the relevant set of state of energy figures within Appendix B). The penalty assigned to individuals resulting in insufficient battery SoE at any point for any vehicle is therefore assumed to be sufficient to ensure that no such individual is selected at the end of the model run.

The average time spent on each segment for all vehicles passing through that segment is given in Appendix C, Figure 121, Appendix C for all those segments with an average time longer than 5 minutes. These are generally cost-efficient locations for electrification since smaller stretches of electrification would service vehicles for longer periods of time. Unsurprisingly, common end stations for trips make up these segments.

It is noteworthy that the algorithm has not placed electrification at these segments. This is discussed further in section 6.2.

Figure 122, Appendix C shows the average battery SoE by segment for all vehicles passing through that segment at the time of traversal. The battery SoE is generally lower between Landau and Pirmasens-Nord hinting at why the longest single stretch of proposed electrification is suggested here. This figure also exemplifies the difference that can exist between vehicles going in different directions on parallel tracks. Near Landstuhl, the average battery SoE varies greatly between the parallel tracks because the vehicles coming from Kaiserslautern have been under catenary whilst vehicles coming from Kusel/Landstuhl have not.

Due to the complicated combinatorial nature of problem, it is difficult to ascertain whether an unnecessarily high quantity of electrification infrastructure is suggested by the optimization approach. Many vehicles in the circulation (for example vehicle 156632701 in Appendix B, Figure 18) are capable of operating as BEMU (as per the definition within this thesis) without any additional electrification infrastructure. Existing infrastructure is sufficient to supply the necessary energy. This is the case for 33% of the vehicles in the circulation. Many vehicles that partly move underneath suggested new catenary do not need it to have access to sufficient energy to complete their circulation. Vehicle 156632716 is an example of this; it is repeatedly charging from new electrification near Kusel even though the new electrification is not required for it to maintain a steady battery SoE through its circulation (refer to Figure 31). However, even though many vehicles do not need the new electrification to maintain battery SoE, some vehicles in the circulation require the catenary to stay above the minimum battery SoE threshold. For the example near Kusel, vehicle 156632712 heavily relies on this particular stretch of electrification. Generally, for vehicles charging and relying on the establishment of new electrification infrastructure, this does not leave them with a surplus of time underneath new catenary relative to the charging time required to maintain a sufficient battery SoE. This indicates that the algorithm has not placed superfluous infrastructure given the problem definition and constraints provided in the thesis. The fact that almost no improvements could be identified for most of the model run strengthens this assumption. However, vehicles do often end their circulation with a lower battery SoE than the initial full charge of 400 kWh (for example vehicle 156632705, Appendix B, Figure 21). This is due to the fact that no further restrictions have been set for the battery SoE at the end of the circulation.

5.3 Sensitivity Analysis

For all model runs within the sensitivity analysis, no vehicle dips below the minimum SoE threshold of 50 kWh. As such, they are all viable solutions to the optimization problem. This section explores how the output of the algorithm responds to changes in the examined parameters and to which the extent the algorithm gravitates towards local minima.

5.3.1 Responses to Parameter Alterations

In most cases, changes to the maximum battery SoE and the cost of substations impact the algorithm output as expected. Refer to Appendix B for all relevant maps and vehicle SoE figures. Figure 14 summarizes the length of catenary and the number of substations suggested in each of the 18 model runs.

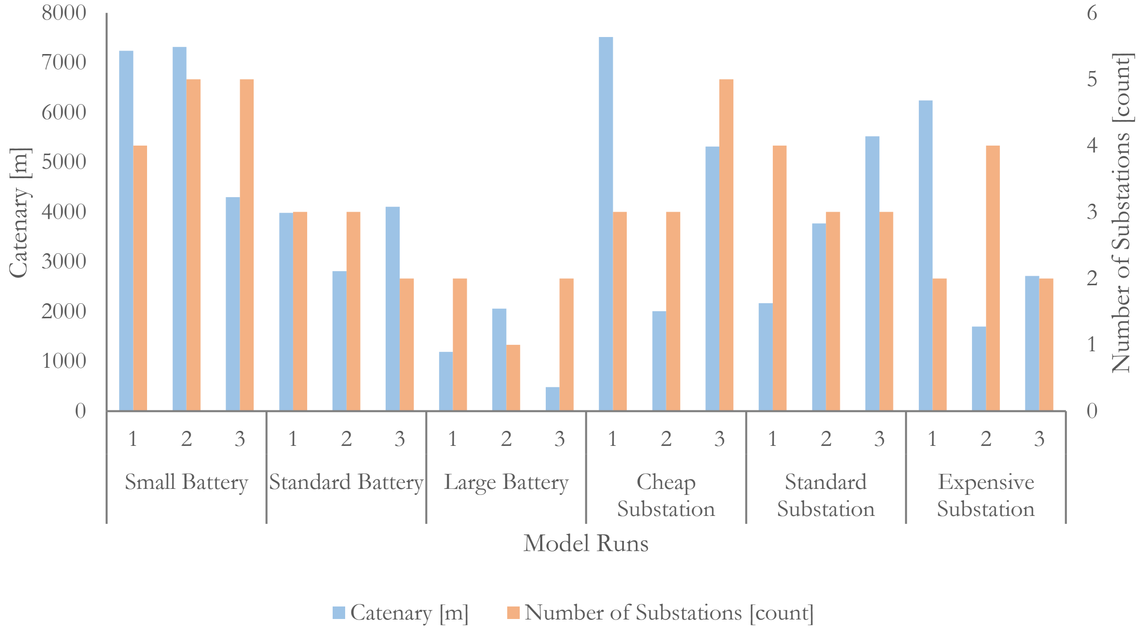

Figure 14: The quantity of catenary and number of substations proposed by all model runs in the sensitivity analysis.

The model runs experimenting with the maximum battery capacity simultaneously explores changes to all parameters in the train vehicle affecting consumption or charging of energy. In line with expectations, less new infrastructure is suggested for the reduced study region when the battery capacity increases. While the small battery model runs places 4.5-7 km of catenary split between an additional 4-5 substations, the large battery model runs only place 1.5-3 km of catenary serviced by an additional 1-2 substations. Especially the battery SoE profile for vehicle 156632707 changes under presumptions of a higher battery capacity. Here, it becomes possible for the vehicle to sustain the first two trips of its circulation entirely without catenary eliminating the need for catenary between Landau and Pirmasens in the model run "large battery 2" (refer to Figure 75 and Figure 76). The changes to battery capacity explored in this thesis had a large impact on the model results.

For the alterations in the price of substations, the impact is less clear. While variation within the identical model run groups combined with the small sample size does not allow for a conclusive image, the average amount of substations does go down as the substation cost increases (7.6 for "cheap" substations, 7.3 for "standard" substations, and 5.3 for "expensive" substations). Model run "expensive substation 1", for example, extends existing electrification west of Saarbrücken Hbf. instead of suggestion a new catenary island (refer to Appendix B, Figure 107). This saves the cost of the substation. The spatial distribution of the suggested electrification varies between identical model runs but it is especially high for model runs assuming a cheaper substation cost (refer to Figure 83, Figure 87 and Figure 91 in Appendix B). Some model runs suggest new electrification between Landau and Pirmasens, some between Kaiserslautern-Pirmasens, some both. When substations matter less, the electrification can be placed more flexibly. This leads to higher variance in the problem solutions.

5.3.2 Convergence on Local Minima

There is a clear tendency for the algorithm optimization approach to seek local minima. Each model run output is a possible problem solution, but it is only one of many possible electrification configurations – some of which are more optimal than others. It is not viable to determine the statistical variation given the limited sample size. Instead, the range of total infrastructure costs in each identical set of three model runs is given on Figure 15. In addition to the variation in the length of catenary and the number of substations symbolizing the spatial distribution of catenary, the total cost of the new infrastructure also varies between model runs with identical input parameters. For the three model runs conducted in each set, the total infrastructure cost varies plus or minus by approximately €5 million. The variation is especially large within the expensive substation group. This is due to model run "expensive substation 3" that manages to converge on a solution requiring only two additional substations without requiring more catenary length than the other problem solutions. The suggested placement of electrification also varies between model runs even in cases where the infrastructure costs do not. If a catenary island is required between two stations, the algorithm has no particular reason to place it at one location in favor of another. This produces variability in the exact spatial distribution between model runs not because of local minima but because of codified indifference. However, even despite this, there is a clear tendency for the optimization approach to gravitate towards local minima.

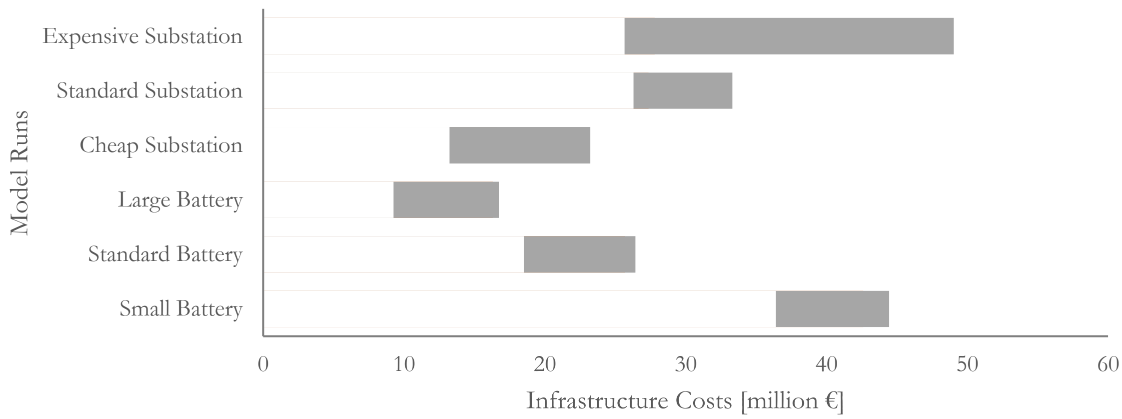

Figure 15: The range of estimated total additional infrastructure costs for the optimized distribution of new charging infrastructure proposed by the model runs in the sensitivity analysis. Each model run type contains three model runs with the same input parameters.

6. Discussion

This section discusses the results, performance, and possible alterations/additions to the thesis methodology. I compare the results of the model run “entire Pfalznetz” to that of an existing BEMU electrification plan serving as a real-world point of comparison for the thesis results. Hereafter, the section reflects on the shortcomings of the thesis methodology and ways in which several aspects could be improved. Lastly, further research opportunities are discussed in the light of the methodology developed within this thesis.

6.1 The Existing Pfalznetz Electrification Plan

A cooperation between the local passenger rail association for Rheinland-Pfalz Süd, the Regional Ministry for Climate Protection, Environment, Energy and Mobility, DB Energy AG, DB Regio AG, and DB Netz AG has initiated a plan for the full electrification of the Pfalznetz network necessary to allow for the replacement of all DMUs with BEMUs. The infrastructure plan, that is scheduled to start gradual phase in from December 2025, contains a total of 14.8 km of additional electrification distributed near the stations Kusel, Kaiserslautern, Lauterecken-Grumbach, Pirmasens-Nord, Landau, and Winden [49] (refer to Figure 16). This recent electrification plan offers an excellent opportunity for real-world comparison between the results of this thesis and more comprehensive plans drawn through different methodologies.

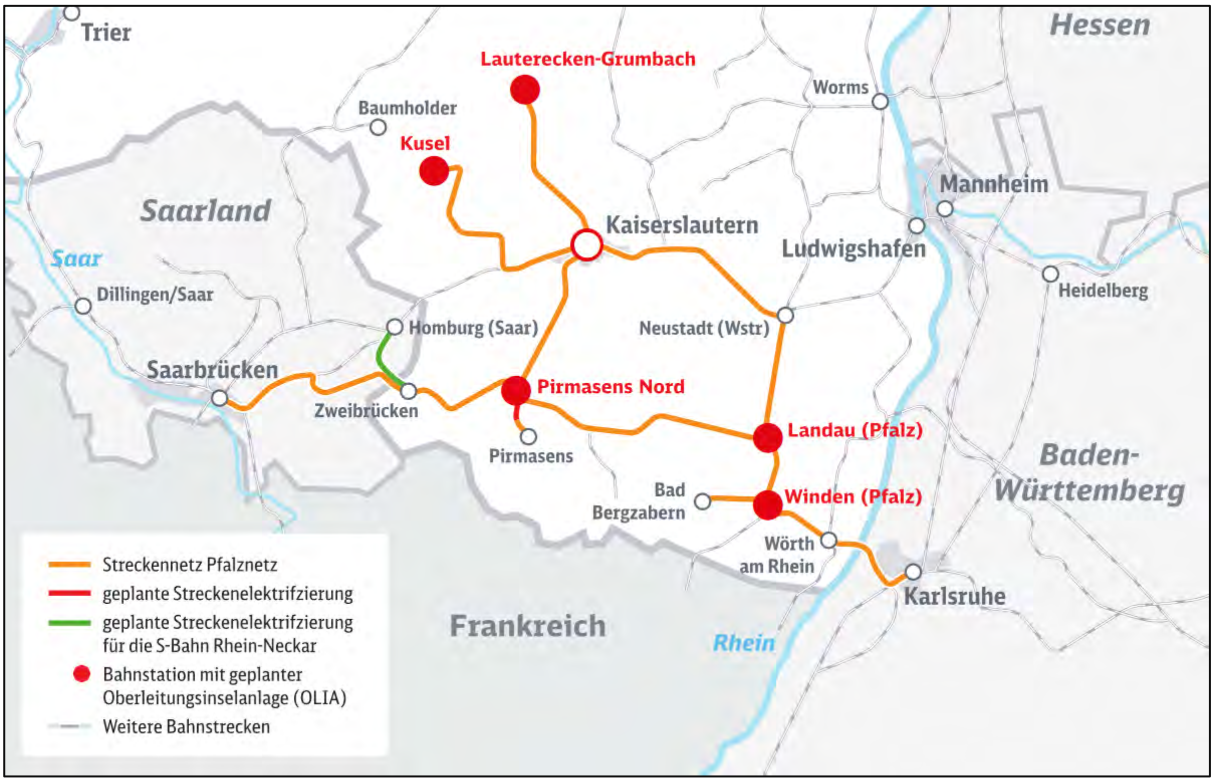

Figure 16: Map showing the planned electrification for Pfalznetz within the Pfalznetz BEMU electrification plan [49]. Translation of legend German to English: Streckennetz Pfalznetz = Rail network Pfalznetz, geplante Streckenelektrifizierung = Planned rail network electrification, geplante Streckenelektrifizierung für die S-Bahn Rhein-Neckar = Planned rail network electrification for the S-Bahn Rhein-Neckar, Bahnstation mit geplanter Oberleitungsinselage = Railway station with planned catenary island, Weitere Bahnstrecken = Other railway lines.

If the cost assessment of the Pfalznetz electrification plan were to follow the same rules as the infrastructure cost estimates employed within this thesis, the electrified length of 14.8 km serviced by 5 additional substations would amount to €49.1 million. This is 31% less than the estimated cost of the electrification distribution proposed by the model run “entire Pfalznetz”.

The spatial distribution between the two suggested configurations also varies greatly. The existing electrification plan places new electrification at stations aligning with the end stations for many trips conducted in the circulation plan. This distribution aligns better with the average time spent at segments visualized on Figure 121, Appendix C than the distribution proposed by the model run of the thesis. As such, given both the lower cost and more reasonable placement of additional infrastructure at efficient locations, the output of the methodology of this thesis appears distant from a real-world scenario.

There are several reasons why this may be the case. While it is not known exactly how the electrification plan has been developed, it is fair to assume that this development included a different methodology to that proposed in this thesis, and that this method would have had access to a much wider and more expansive dataset. With a better understanding of the vehicle circulation and the network itself, a better electrification plan can be developed. This increased understanding presumably also derives from site-specific infrastructure assessments and cooperation with vehicle construction partners to design the infrastructure in accordance with vehicle specifications. Additionally, it is possible that changes to the circulation plan itself is proposed in conjunction with the electrification strategy to align the movement of the vehicles with new infrastructure requirements. This is a concept unexplored within the thesis scope. However, with amendments to the methodology discussed in section 6.2, I believe it is likely that the model output would align better with the proposed electrification plan much more closely. This is especially the case for amendments related to battery SoE at the end of a vehicle circulation discussed in section 6.2.2.

6.2 Methodological Reflections

The development of the methodological approach to optimize the spatial distribution of electrification infrastructure for BEMUs is main contribution of the thesis to the research at large. However, there are many avenues by which the approach could be improved or altered to achieve a better result. In light of that, this section discusses the quality of the input data, the efficiency and design of the genetic algorithm, and its tendency to approach local minima.

6.2.1 Quality and Availability of Input Data

The optimization approach employed in this thesis builds on a limited dataset together with necessary assumptions drawn without confirmation backed up by data. Consequently, this limits the quality of the optimized output proportional to the quality and quantity of input.

Several assumptions were drawn relating to the nature of network. All electrification data is derived from OSM which, while it does maintain a mostly high quality of data, is not an official source for the location of built electrification present in the study area. Furthermore, and very importantly, the actual spatial movement of the vehicles through their trips is unknown and entirely presumed based on timetable information. Which tracks the vehicles use is paramount to the optimization problem and the fact that this is presumed is a key weakness in the input data of the thesis. This is amplified by the fact that the Pfalznetz circulation plan and the GTFS dataset do not match in their entirety. Some vehicles had to be excluded and others had their circulation reduced to the consecutive elements that could be matched with the GTFS dataset. As such, the circulation used is not an accurate representation of reality but rather a modified version of an actual vehicle circulation plan.

Core input parameters for the optimization are also assumed and should be re-evaluated (and/or expanded) to make the algorithm more suitable for real-world use. The selected train vehicle is a generic assumption of how such a vehicle might function. While it works for testing purposes, it does not correspond with any existing vehicle that could operate in the Pfalznetz region. More prudent research on this area or cooperation with a vehicle manufacturer could significantly improve this aspect of the methodology.

This is also the case for infrastructure costs that are highly simplified in their current implementation. Infrastructure construction depends on a range of factors among other elements involving land acquisition, material costs, construction site conditions, soil pedology, and access to transportation infrastructure. Specifically for rail electrification, the current extent, characteristics, and capacities of the energy grid have a large influence on the most suitable locations for electrification infrastructure. None of these considerations are included in the thesis methodology. Were they to be included, this could markedly alter the optimized outputs.

6.2.2 The Efficiency and Functioning of the Genetic Algorithm

Processing time is the main limitation on the scale on which the optimization approach can be implemented. With better processing performance it would be possible to increase population sizes and the volatility of the mutation and crossover procedures. This would allow for the algorithm to explore a greater extent of the problem space and thereby make it more likely to identify better solutions. So, too, would a sensitivity analysis be able to be conducted on a statistically useful quantity of model runs to achieve a more conclusive understanding of the model parameters and output variance. It is also relevant to improve processing times in unison with the expansion of the algorithm functionalities. With increased complexity often comes increased processing time, and the algorithm's already high computational requirements necessitates further improvement for it to evolve in its problem scope.

There are several ways this could be achieved. Nearly the entirety of the computation time is related to the calculation of individual fitness. The code for this is based in pandas dataframes [50] which, while intuitive to read and code, is not the fastest for data processing. An approachable alteration could be to work with numpy arrays instead [51], but more advanced data handling libraries and techniques could improve processing time even further. The fitness function also uses .iterrows() [52] specifically when calculating battery SoE. This is a slow processing approach for which a replacement should be identified. Iterrows have been avoided in all other aspects of the algorithm, including inside the fitness function. There may also be optimizations possible in how infrastructure costs are calculated – especially relating to the buffers generated and dissolved around each segment to estimate substation count. This is a costly operation that should be optimized further.

An easier to implement alteration to the methodology that would reduce processing time relates to the fact that many vehicles in the circulation do not need additional electrification to operate. For the purposes of this study, they could have been excluded entirely from the optimization thereby reducing the size of the examined circulation plan by 33%. This would translate almost directly to a 33% reduction in processing time for the "entire Pfalznetz" model run. A simple pre-calculation of the battery SoE for all vehicles without additional electrification could have identified the vehicles suitable for exclusion.

As a final note on the functioning of the genetic algorithm, it is relevant to re-evaluate how vehicles are treated at the end of their circulation. Currently, there is no implemented threshold on the battery SoE a vehicle must have at the end of the circulation. Combined with the fact that many vehicles end their circulation on non-electrified stations, the suggested electrification configuration would not allow the vehicles to repeat their circulation the following day. This could be remedied by simply requiring vehicles to end their circulation with full battery SoE whilst allowing for this to be completed overnight if electrification is present. This way, the algorithm would be much more incentivized to place electrification at end stations and the output likely much more similar to that of the real-world Pfalznetz electrification plan.

6.2.3 Avoiding Local Minima

As shown in the results of the sensitivity analysis (refer to section 5.3.2), the current design of the genetic algorithm has a strong inclination to seek local minima. The algorithm output consistently varies when provided unchanging input. This is an issue for the optimization tasks aiming to identify not just a possible electrification configuration, but the most optimal configuration of all possible options.

Tuning the parameters and functioning of the mutation and crossover processes is key to solving the issue of local minima convergence. While the current implementation of fine-grained mutation detailed in section 4.5.4 allows for the algorithm to fine tune a distribution of electrification, it is also the reason the algorithm struggles with convergence on local minima. Given that only small incremental changes can occur from one generation to the next, it is not possible for the algorithm to explore scenarios where an entire island of electrification is placed in a completely different location. Instead, the algorithm focuses on optimizing a draft configuration of electrification that is identified at random early in the model run. Solutions for this could include an even further complex implementation of mutations, where the mutation probability could vary across generations and individuals. Means through which to change the electrification state of contiguous segments of track simultaneously should also be explored.

The variation in the precise placement of electrification between identical model runs is also largely due to the fact that the algorithm is fed no parameters for which to decide on one location over another. If a stretch between two stations needs some amount of electrification to service the vehicle circulation, the precise placement of this is largely left up to chance. The most fitting method to remedy this issue would be to increase the complexity of input data and algorithm design including more relevant parameters (like those discussed in section 6.2.1). This would provide the algorithm reasons with which to evaluate certain locations as more preferable to others.

6.3 Further Research

In addition to the changes and alterations suggested in section 6.2, there are broader expansions of the methodological concept that could be worth exploring in further research. As an optimization tool, the genetic algorithm is inherently flexible in its application. The thesis draws on that quality to build a tool with flexible input parameters and data. This could be expanded much further without changing the built code structure.

One interesting possible path of expanding the methodology would be to include evaluations of required energy draw from the perspective of the energy grid. Since the methodology already computes energy transferred at each existing and suggested segment of electrification on the network in space and time, this information could be harnessed in combination with understandings of the energy grid to produce valuable insights. The energy transfer profile at each segment could be evaluated in parallel with the battery SoE of the vehicles. Understandings of the limitations and possibilities of the energy grid could be included to limit or incentivize the algorithm to electrify certain segments in favor of others. Energy pricing could also be included, and these could be dependent on both space and time. Both implementations would require only minor expansions of existing code. Due to limitations in scope, this was excluded from the thesis methodology.

A common approach in related research is to include the battery sizing within the optimization approach [33], [36]. This too could be implemented within this methodology. This way, the methodology would better recognize the interplay between optimizing charging infrastructure and the vehicles using the infrastructure. When none are determined in advance, as is the case for emerging BEMU vehicles, the possibility to optimize them as a combinatorial pair has greater potential than optimizing their designs separately.

Lastly, further research could attempt to reconsider stations themselves within the genetic algorithm optimization. The current implementation does not explicitly account for stations as a distinct typology. Only track segments exist, and for some of these segments vehicles may spend large amount of stationary time. These segments are present at stations. This implementation is simplistic and consistent but suffers from the lack of precise information on which tracks vehicles use at the stations and at which precise point on the tracks the vehicles remain stationary. This results in inconsistencies between vehicle circulations at which segments a station exists. Presuming the source data could not be improved, an alternative data structure could better accommodate the uncertainty. For the purposes of the optimization task, it may be preferable to store stations are singular points where all relevant segments converge. This would standardize the positioning of stations and allow for the algorithm to better identify when several vehicles (or the same vehicle throughout its circulation) pass through or stop at the same station.

7. Conclusion

This thesis presents a methodological approach to optimize the spatial allocation of charging infrastructure to allow for the operation of BEMUs instead of DMUs on regional passenger rail. The approach is comprised of a genetic algorithm optimization relying upon the processing and formatting of input data from OSM, a DTM, route geometries, GTFS data, and simulated vehicle power at wheel. The methodology successfully optimizes the spatial distribution of charging infrastructure given its limited conceptualizations of the real-world scenario for the model runs conducted within the scope of the thesis. It does so considering single and parallel tracks, existing electrification, multiple vehicles circulation simultaneously, and with an understanding of infrastructure costs as both by electrified track length and by independent contiguous islands of catenary. A sensitivity analysis indicates that the model behaves expectedly to alterations of key parameters relating to the circulation vehicle and infrastructure costs. For Pfalznetz, the model suggests a spatial distribution of charging infrastructure with a combined cost of €72.4 million of which substation costs make up €64 million and catenary €8.4 million.

However, the approach suffers from several shortcomings relating to the quality and availability of input data, the efficiency and functioning of the algorithm and the approaching of local instead of global minima. Many of these issues could be improved in future iterations of the approach. Some minor code alternations, like properly accounting for overnight charging, may substantially alter the model output of the algorithm further aligning its output with real-world electrification plans. The approaching of local minima could also be reduced with alternations and expansions of how mutation is handled within the genetic algorithm. Lastly, the methodology could be expanded in various directions to consider larger variety of aspects relating to infrastructure placement or to consider currently unexplored elements. A promising further inclusion could be integrating an evaluation from the perspective of the power grid of the energy draw that results from the electrification configuration and the vehicle circulation.

8. References

[1] Intergovernmental Panel On Climate Change (IPCC), *Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, 1st ed. Cambridge University Press, 2023. doi: 10.1017/9781009325844.

[2] Bundesministerium für Wirtshaft und Klimaschutz, "Deutsche Klimaschutzpolitik." Accessed: Sep. 29, 2022. [Online]. Available: https://www.bmwk.de/Redaktion/DE/Artikel/Industrie/klimaschutz-deutsche-klimaschutzpolitik.html

[3] M. Lambrecht, "Klimaschutz im Verkehr," Umweltbundesamt. Accessed: Sep. 29, 2022. [Online]. Available: https://www.umweltbundesamt.de/themen/verkehr-laerm/klimaschutz-im-verkehr

[4] Deutsche Bahn, "Infrastrukturzustands- und -entwicklungsbericht 2020." Deutsche Bahn, 2020.

[5] "Means of transport and infrastructure," Federal Statistical Office. Accessed: Sep. 22, 2023. [Online]. Available: https://www.destatis.de/EN/Themes/Economic-Sectors-Enterprises/Transport/Enterprises-Infrastructure-Vehicle-Stock/Tables/railway-infrastructure.html

[6] E. Holden, D. Banister, S. Gössling, G. Gilpin, and K. Linnerud, "Grand Narratives for sustainable mobility: A conceptual review," *Energy Res. Soc. Sci.*, vol. 65, p. 101454, Jul. 2020, doi: 10.1016/j.erss.2020.101454.

[7] T.-V. Nguyen, J. Schnidrig, and F. Marechal, "An analysis of the impacts of green mobility strategies and technologies on different European energy systems," in *34th International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems (ECOS 2021)*, Taormina, Italy: ECOS 2021 Program Organizers, 2022, pp. 941–959. doi: 10.52202/062738-0084.

[8] N. D. Popovich, D. Rajagopal, E. Tasar, and A. Phadke, "Economic, environmental and grid-resilience benefits of converting diesel trains to battery-electric," *Nat. Energy*, vol. 6, no. 11, Art. no. 11, Nov. 2021, doi: 10.1038/s41560-021-00915-5.

[9] J. J. Mwambeleko, K. Somsai, and T. Kulworawanichpong, "The potential of battery electric multiple units to replace diesel commuter trains and reduce fuel cost," in *2016 IEEE/SICE International Symposium on System Integration (SII)*, Dec. 2016, pp. 25–30. doi: 10.1109/SII.2016.7843970.

[10] C. Streuling, J. Pagenkopf, M. Schenker, and K. Lakeit, "Techno-Economic Assessment of Battery Electric Trains and Recharging Infrastructure Alternatives Integrating Adjacent Renewable Energy Sources," *Sustainability*, vol. 13, no. 15, p. 8234, Jul. 2021, doi: 10.3390/su13158234.

[11] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," *Multimed. Tools Appl.*, vol. 80, no. 5, pp. 8091–8126, Feb. 2021, doi: 10.1007/s11042-020-10139-6.

[12] S. Mirjalili, *Evolutionary Algorithms and Neural Networks*, vol. 780. in Studies in Computational Intelligence, vol. 780. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-319-93025-1.

[13] Joon-Yong Lee, Min-Soeng Kim, Cheol-Taek Kim, and Ju-Jang Lee, "Study on encoding schemes in compact genetic algorithm for the continuous numerical problems," *SICE Annu. Conf. 2007*, pp. 2694–2699, Sep. 2007, doi: 10.1109/SICE.2007.4421447.

[14] K. Jebari, "Selection Methods for Genetic Algorithms," *Int. J. Emerg. Sci.*, vol. 3, pp. 333–344, Dec. 2013.

[15] G. K. Soon, T. T. Guan, C. K. On, R. Alfred, and P. Anthony, "A comparison on the performance of crossover techniques in video game," in *2013 IEEE International Conference on Control System, Computing and Engineering*, Nov. 2013, pp. 493–498. doi: 10.1109/ICCSCE.2013.6720015.

[16] M. Albadr, S. Tiun, M. Ayob, and F. Al-Dhief, "Genetic Algorithm Based on Natural Selection Theory for Optimization Problems," *Symmetry*, vol. 12, pp. 1–31, Oct. 2020, doi: 10.3390/sym12111758.

[17] A. Lambora, K. Gupta, and K. Chopra, "Genetic Algorithm- A Literature Review," in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, Feb. 2019, pp. 380–384. doi: 10.1109/COMITCon.2019.8862255.

[18] M. Kumar, M. Husain, N. Upreti, and D. Gupta, "Genetic Algorithm: Review and Application," *SSRN Electron. J.*, 2010, doi: 10.2139/ssrn.3529843.

[19] L. Haldurai, T. Madhubala, and R. Rajalakshmi, "A Study on Genetic Algorithm and its Applications," *Int. J. Comput. Sci. Eng.*, vol. 4, p. 6, 2016.

[20] S. S. Chouhan, A. Kaul, and U. P. Singh, "Soft computing approaches for image segmentation: a survey," *Multimed. Tools Appl.*, vol. 77, no. 21, pp. 28483–28537, Nov. 2018.

[21] A. Khan *et al.*, "Color image segmentation using genetic algorithm with aggregation-based clustering validity index (CVI)," *Signal Image Video Process.*, vol. 13, no. 5, pp. 833–841, Jul. 2019, doi: 10.1007/s11760-019-01419-2.
[22] N. Alaoui, A. B. H. Adamou-Mitiche, and L. Mitiche, "Effective hybrid genetic algorithm for removing salt and pepper noise," *IET Image Process.*, vol. 14, no. 2, pp. 289–296, 2020, doi: 10.1049/iet-ipr.2019.0566.
[23] L. Tang, L. Tian, and B. L. Steward, "Color Image Segmentation with Genetic Algorithm for In-field Weed Sensing," *Trans. ASAE*, vol. 43, no. 4, pp. 1019–1027, 2000, doi: 10.13031/2013.2970.
[24] A. Peerlinck, J. Sheppard, J. Pastorino, and B. Maxwell, "Optimal Design of Experiments for Precision Agriculture Using a Genetic Algorithm," in *2019 IEEE Congress on Evolutionary Computation (CEC)*, Wellington, New Zealand: IEEE, Jun. 2019, pp. 1838–1845. doi: 10.1109/CEC.2019.8790267.
[25] N. J. Camargo, G. E. Meyer, and D. D. Jones, "Individual leaf extractions from young canopy images using Gustafson-Kessel clustering and a genetic algorith," *Comput. Electron. Agric.*, vol. 51, no. 1–2, pp. 66–85, Apr. 2006, doi: 10.1016/j.compag.2005.11.002.
[26] W. Roush, "Welcome to Google Transit: How (and Why) the Search Giant is Remapping Public Transportation," *Community Transp.*, 2012, Accessed: Apr. 19, 2023. [Online]. Available: https://www.semanticscholar.org/paper/Welcome-to-Google-Transit%3A-How-(and-Why)-the-Search-Roush/9ea32cef3241904fee2f11819493a99724c652bc
[27] "GTFS | Static Transit," Google Developers. Accessed: Sep. 28, 2022. [Online]. Available: https://developers.google.com/transit/gtfs/reference
[28] A. Khani, M. Hickman, and H. Noh, "Trip-Based Path Algorithms Using the Transit Network Hierarchy," *Netw. Spat. Econ.*, vol. 15, no. 3, pp. 635–653, Sep. 2015, doi: 10.1007/s11067-014-9249-3.
[29] E. Chondrodima, H. Georgiou, N. Pelekis, and Y. Theodoridis, "Public Transport Arrival Time Prediction Based on GTFS Data," in *Machine Learning, Optimization, and Data Science*, G. Nicosia, V. Ojha, E. La Malfa, G. La Malfa, G. Jansen, P. M. Pardalos, G. Giuffrida, and R. Umeton, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2022, pp. 481–495. doi: 10.1007/978-3-030-95470-3_36.
[30] R. H. M. Pereira, P. R. Andrade, and J. P. B. Vieira, "Exploring the time geography of public transport networks with the gtfs2gps package," *J. Geogr. Syst.*, Dec. 2022, doi: 10.1007/s10109-022-00400-x.
[31] N. Wessel and S. Farber, "On the accuracy of schedule-based GTFS for measuring accessibility," *J. Transp. Land Use*, vol. 12, no. 1, pp. 475–500, 2019.
[32] D. Tzamakos, C. Iliopoulou, and K. Kepaptsoglou, "Electric bus charging station location optimization considering queues," *Int. J. Transp. Sci. Technol.*, Mar. 2022, doi: 10.1016/j.ijtst.2022.02.007.
[33] X. Wang, C. Yuen, N. U. Hassan, N. An, and W. Wu, "Electric Vehicle Charging Station Placement for Urban Public Bus Systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 1, pp. 128–139, Jan. 2017, doi: 10.1109/TITS.2016.2563166.
[34] T. Uslu and O. Kaya, "Location and capacity decisions for electric bus charging stations considering waiting times," *Transp. Res. Part Transp. Environ.*, vol. 90, p. 102645, Jan. 2021, doi: 10.1016/j.trd.2020.102645.
[35] G. Chen, D. Hu, and S. Chien, "Optimizing Battery-Electric-Feeder Service and Wireless Charging Locations With Nested Genetic Algorithm," *IEEE Access*, vol. 8, pp. 67166–67178, 2020, doi: 10.1109/ACCESS.2020.2985168.
[36] G. Chen, D. Hu, S. Chien, L. Guo, and M. Liu, "Optimizing Wireless Charging Locations for Battery Electric Bus Transit with a Genetic Algorithm," *Sustainability*, vol. 12, no. 21, Art. no. 21, Jan. 2020, doi: 10.3390/su12218971.
[37] M. Akbari, M. Brenna, and M. Longo, "Optimal Locating of Electric Vehicle Charging Stations by Application of Genetic Algorithm," *Sustainability*, vol. 10, no. 4, p. 1076, Apr. 2018, doi: 10.3390/su10041076.
[38] X. Yan, C. Duan, X. Chen, and Z. Duan, "Planning of Electric Vehicle charging station based on hierarchic genetic algorithm," in *2014 IEEE Conference and Expo Transportation Electrification Asia-Pacific (ITEC Asia-Pacific)*, Aug. 2014, pp. 1–5. doi: 10.1109/ITEC-AP.2014.6941087.

[39] M. D. Baumeister, M. M. Wazifehdust, M. M. Salih, and M. Zdrallek, "Optimal Catenary Planning of Trolleybus Systems," *ETG-Kongr. 2021*, pp. 197–202, 2021.
[40] D. Roch-Dupré, T. Gonsalves, A. P. Cucala, R. R. Pecharromán, Á. J. López-López, and A. Fernández-Cardador, "Multi-stage optimization of the installation of Energy Storage Systems in railway electrical infrastructures with nature-inspired optimization algorithms," *Eng. Appl. Artif. Intell.*, vol. 104, p. 104370, Sep. 2021, doi: 10.1016/j.engappai.2021.104370.
[41] "BRouter web client." Accessed: Aug. 30, 2023. [Online]. Available: https://brouter.de/brouter-web/#map=5/50.958/9.844/standard
[42] J. Polster, "Konzeptionierung und Implementierung einer Geodatenbank für die Berechnung und Visualisierung des Energiebedarfs von alternativen Antrieben im Schienenpersonennahverkehr," Freie Universität Berlin, Berlin, 2021.
[43] "Geofabrik Download Server." Accessed: Sep. 09, 2023. [Online]. Available: https://download.geofabrik.de/
[44] N. Earth Science Data Systems, "Earthdata," Earthdata. Accessed: Sep. 01, 2023. [Online]. Available: https://www.earthdata.nasa.gov/homepage
[45] M. Schenker and F. Kuhlkamp, "Optimization Model for Operation of Battery Multiple Units on Partly Electrified Railway Lines," in *2021 Sixteenth International Conference on Ecological Vehicles and Renewable Energies (EVER)*, Monte-Carlo, Monaco: IEEE, May 2021, pp. 1–8. doi: 10.1109/EVER52347.2021.9456608.
[46] M. Schenker, T. Schirmer, and H. Dittus, "Application and improvement of a direct method optimization approach for battery electric railway vehicle operation," *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit*, vol. 235, no. 7, pp. 854–865, Aug. 2021, doi: 10.1177/0954409720970002.
[47] Felix-Antoine Fortin and M. D. R. Franc, "DEAP: Evolutionary Algorithms Made Easy," *J. Mach. Learn. Res.*, vol. 13, pp. 2171–2175, 2012.
[48] "Crossover in Genetic Algorithm," GeeksforGeeks. Accessed: Sep. 09, 2023. [Online]. Available: https://www.geeksforgeeks.org/crossover-in-genetic-algorithm/
[49] "DB - battery train Pfalznetz." Accessed: Sep. 17, 2023. [Online]. Available: https://www.akkuzug-pfalznetz.de/
[50] "pandas documentation — pandas 2.1.1 documentation." Accessed: Sep. 27, 2023. [Online]. Available: https://pandas.pydata.org/docs/index.html
[51] "NumPy." Accessed: Sep. 27, 2023. [Online]. Available: https://numpy.org/
[52] "pandas.DataFrame.iterrows — pandas 2.1.0 documentation." Accessed: Sep. 16, 2023. [Online]. Available: https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.iterrows.html

9. Appendix A: Circulation Plan

This appendix contains parts of the transcribed, edited, fixed, and connected Pfalznetz circulation plan used in the model runs. It provides information on when and where each vehicle moves as it traverses through its circulation. Columns only relevant to data processing, and not to data interpretation, are excluded. The exact movement of each trip is stored in the GTFS timetable data and the routed geometries (not provided in the appendix).

Vehicle number	Trip id	Start station	End station	Start time
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]

Vehicle number	Trip id	Start station	End station	Start time
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Vehicle number	Trip id	Start station	End station	Start time
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]

Vehicle number	Trip id	Start station	End station	Start time
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]

Vehicle number	Trip id	Start station	End station	Start time
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	

Vehicle number	Trip id	Start station	End station	Start time
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]

Vehicle number	Trip id	Start station	End station	Start time
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	
[redacted]	[redacted]			[redacted]
[redacted]	[redacted]		[redacted]	

Table 12: Complete formatted and linked (with the GTFS dataset) circulation plan for the subnetwork of Pfalznetz used as input in the thesis methodology.

10. Appendix B: Model Run Results

This appendix contains the main results for all model runs conducted in the thesis. The key results for each model run are visualized in a map showing the proposed new electrification on the railway network accompanied by battery state of energy graphs for each circulating vehicle in the model run.

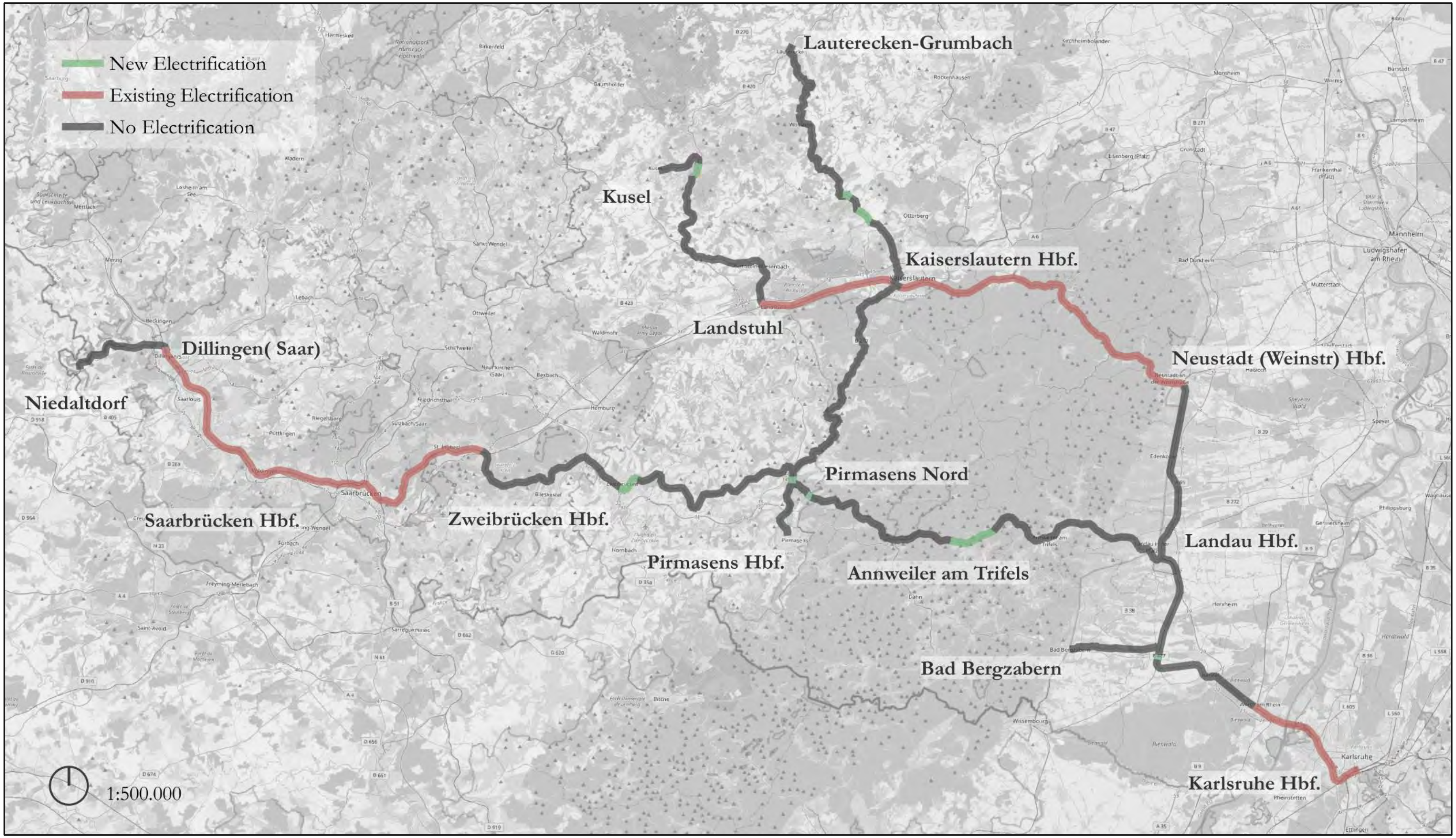

Figure 17: Existing and proposed electrification for model run "entire Pfalznetz". The electrification is given per track even though this is not visible at map scale.

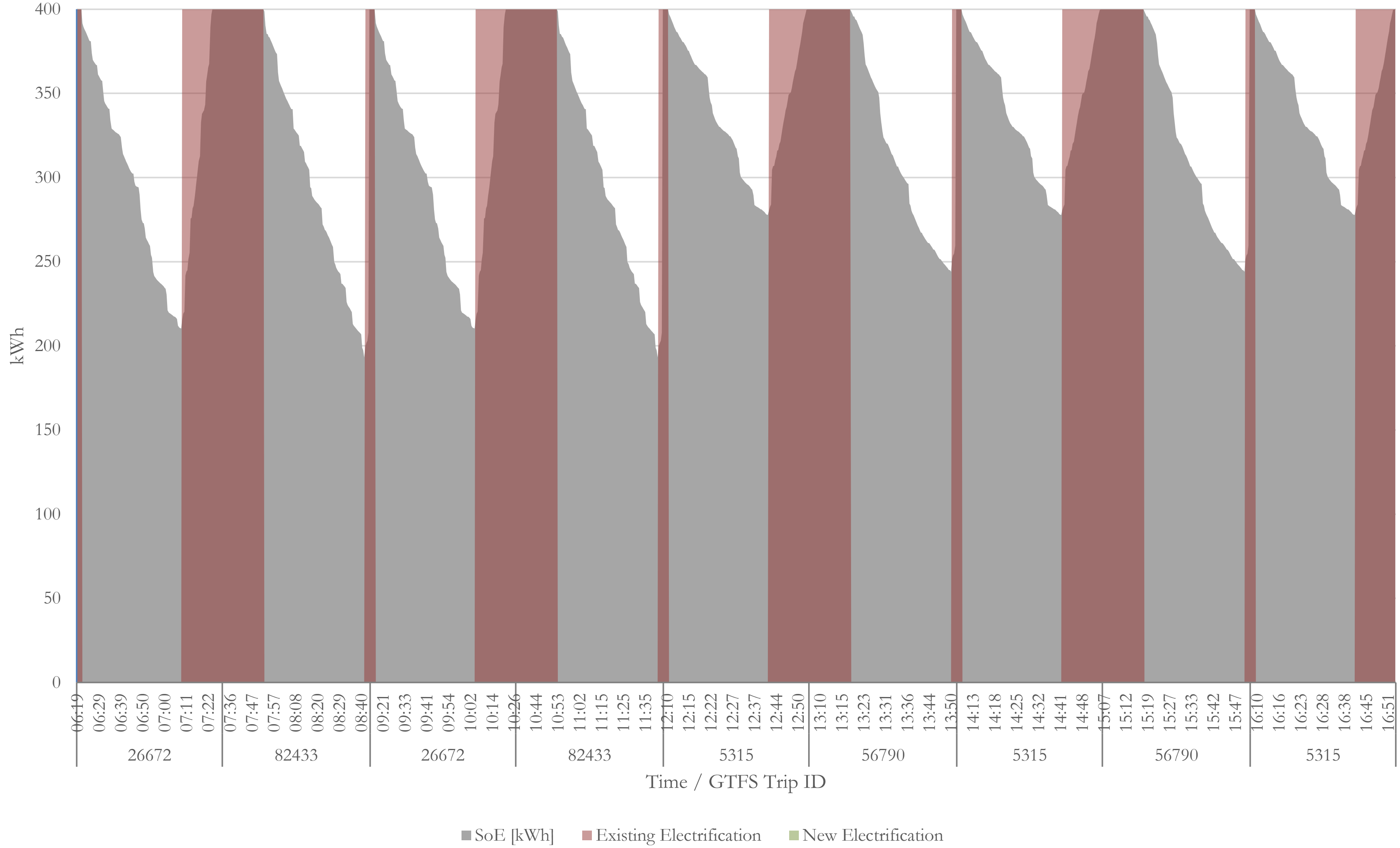

Figure 18: State of energy and the presence of electrification for vehicle 156632701 through its circulation given the distribution of existing and new electrification within the results of model run "entire Pfalznetz.

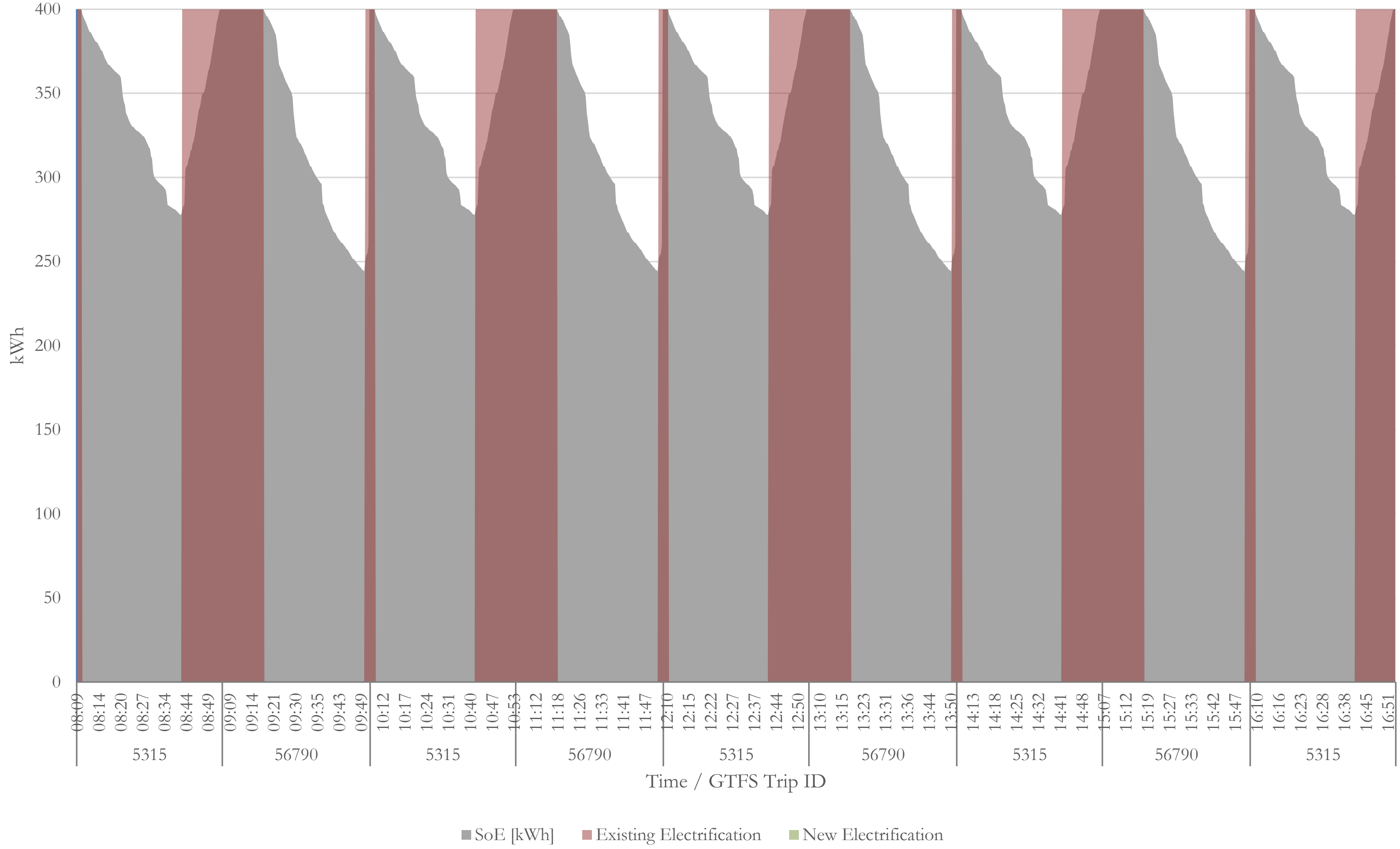

Figure 19: State of energy and the presence of electrification for vehicle 156632702 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

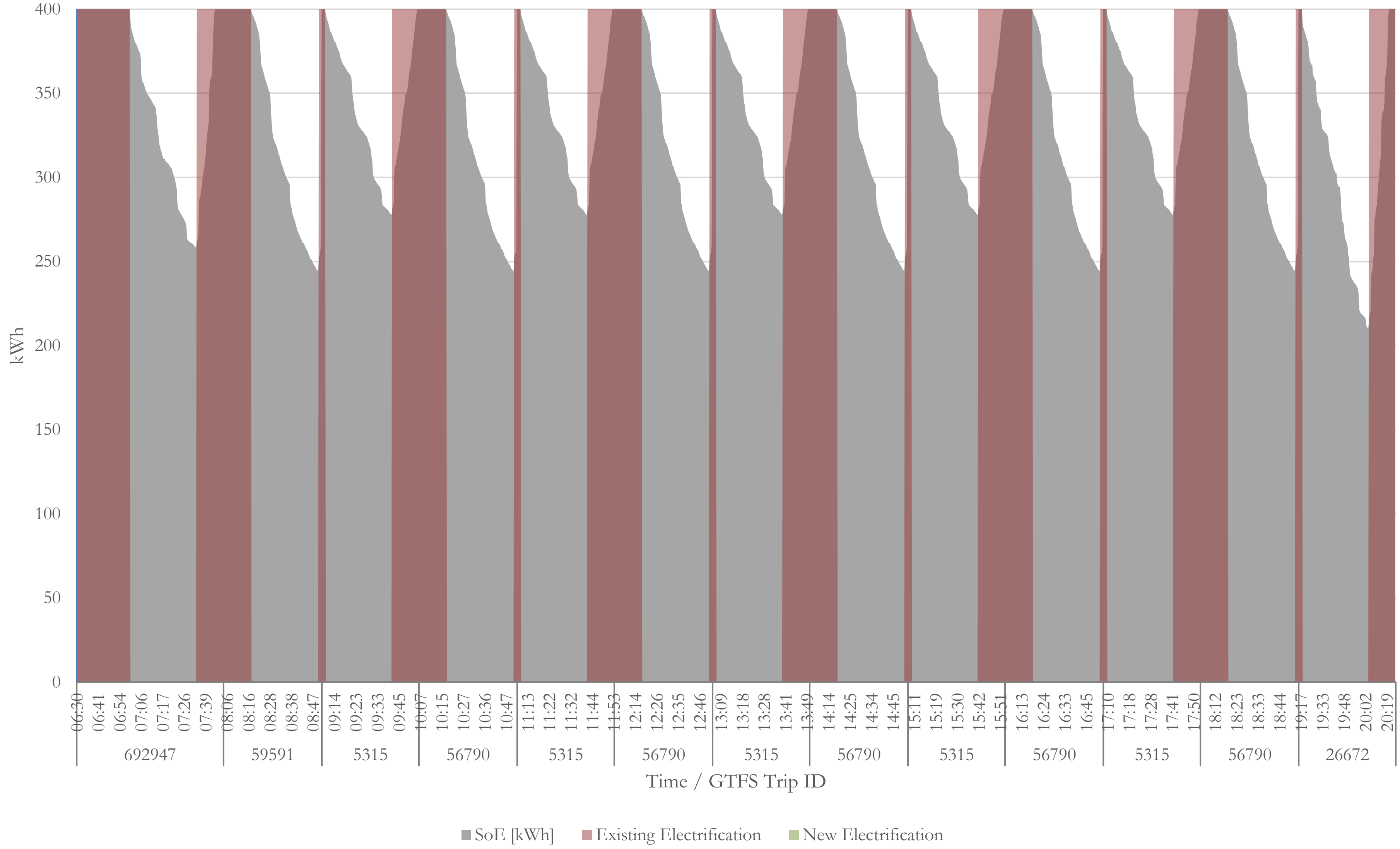

Figure 20: State of energy and the presence of electrification for vehicle 156632703 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

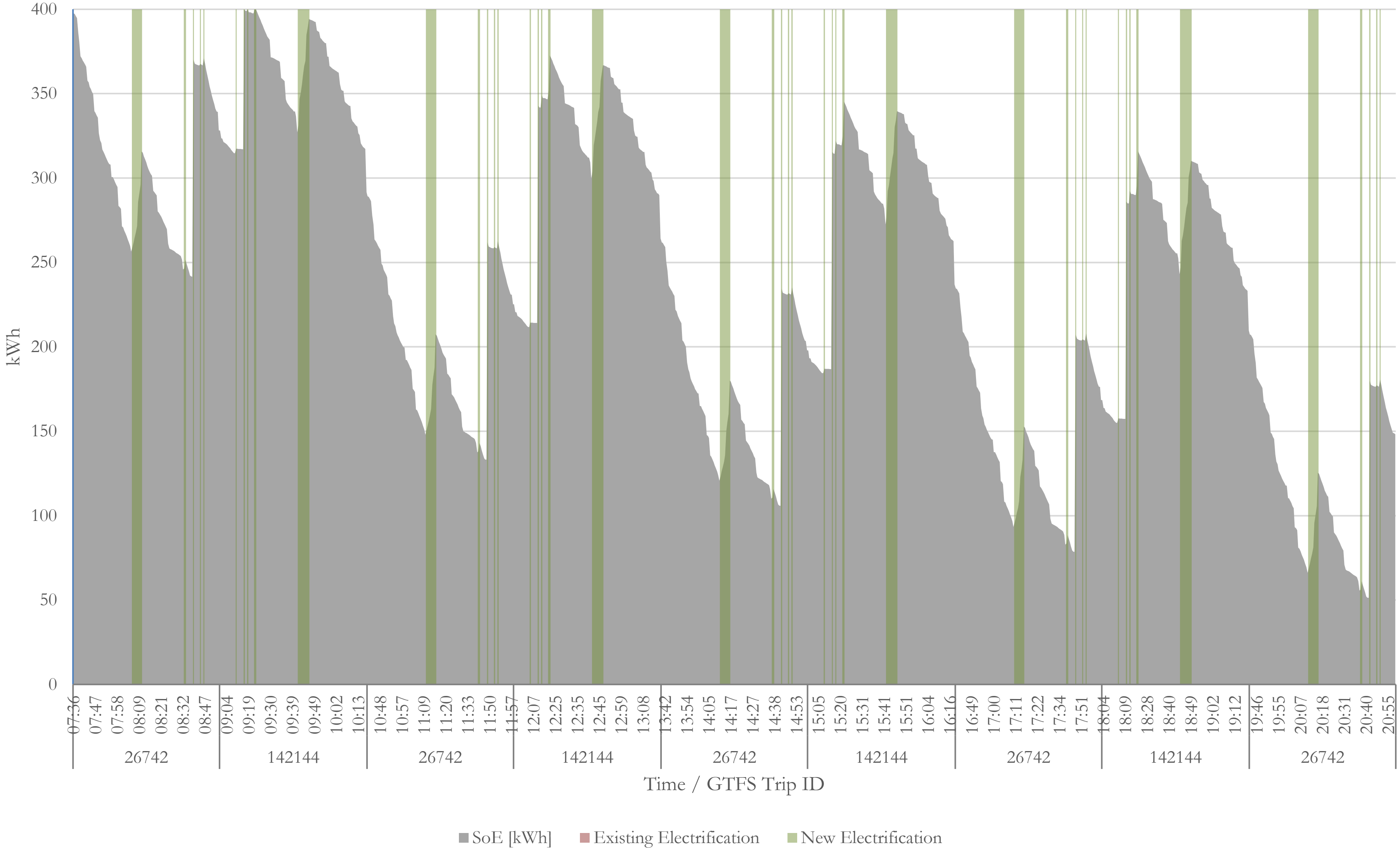

Figure 21: State of energy and the presence of electrification for vehicle 156632705 through its circulation given the distribution of existing and new electrification within the results of model run "entire Pfalznetz.

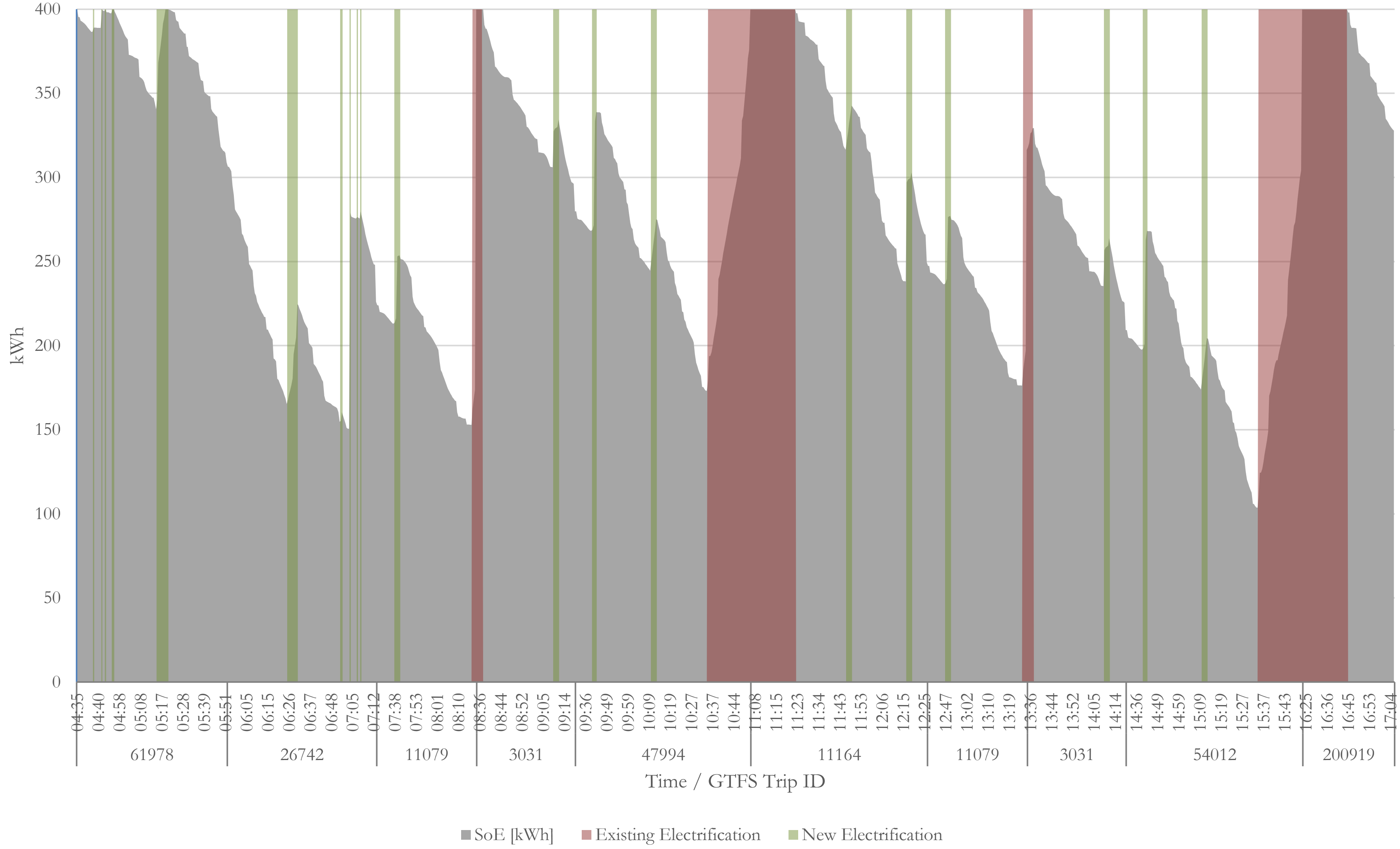

Figure 22: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run "entire Pfalznetz.

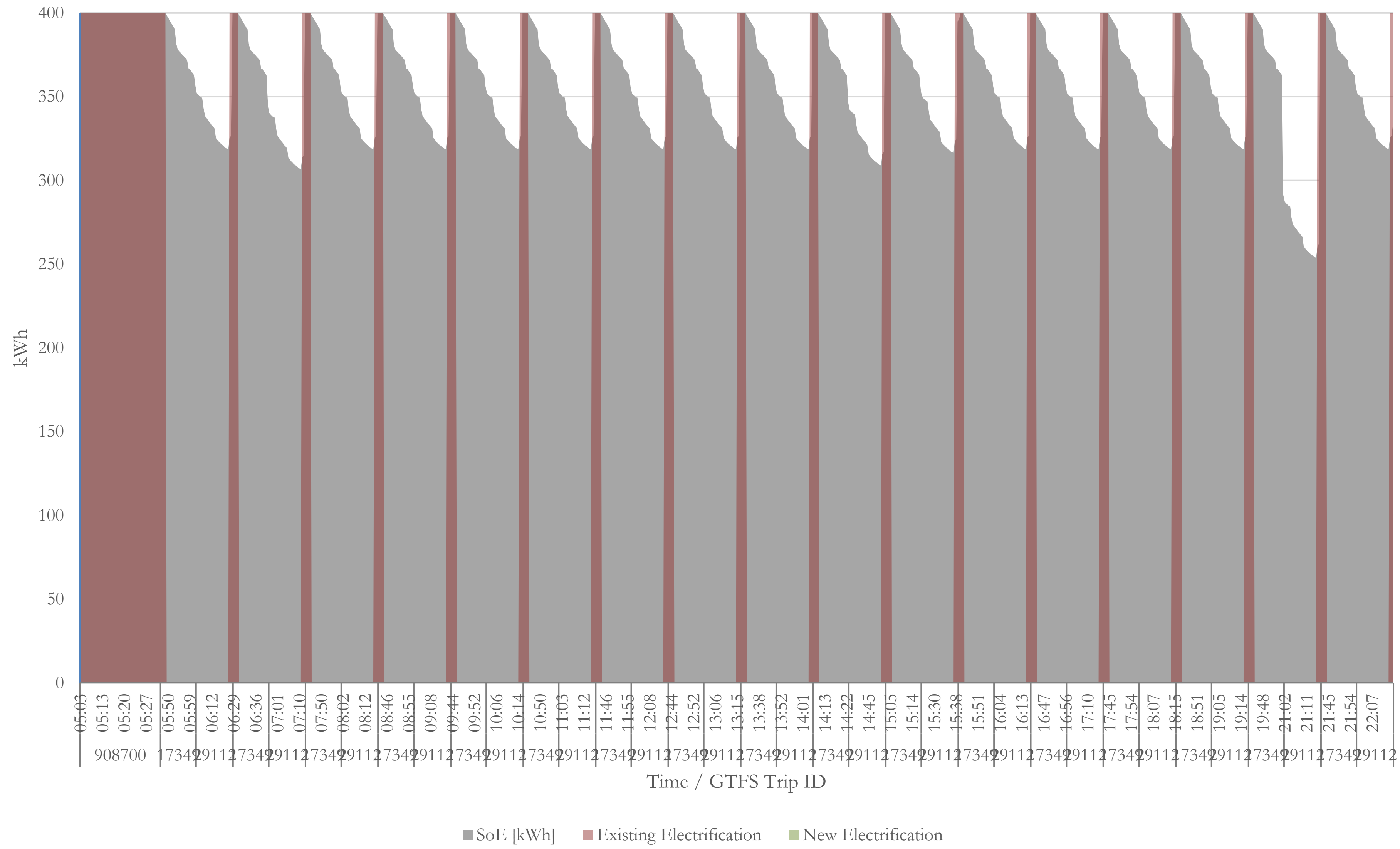

Figure 23: State of energy and the presence of electrification for vehicle 156632708 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

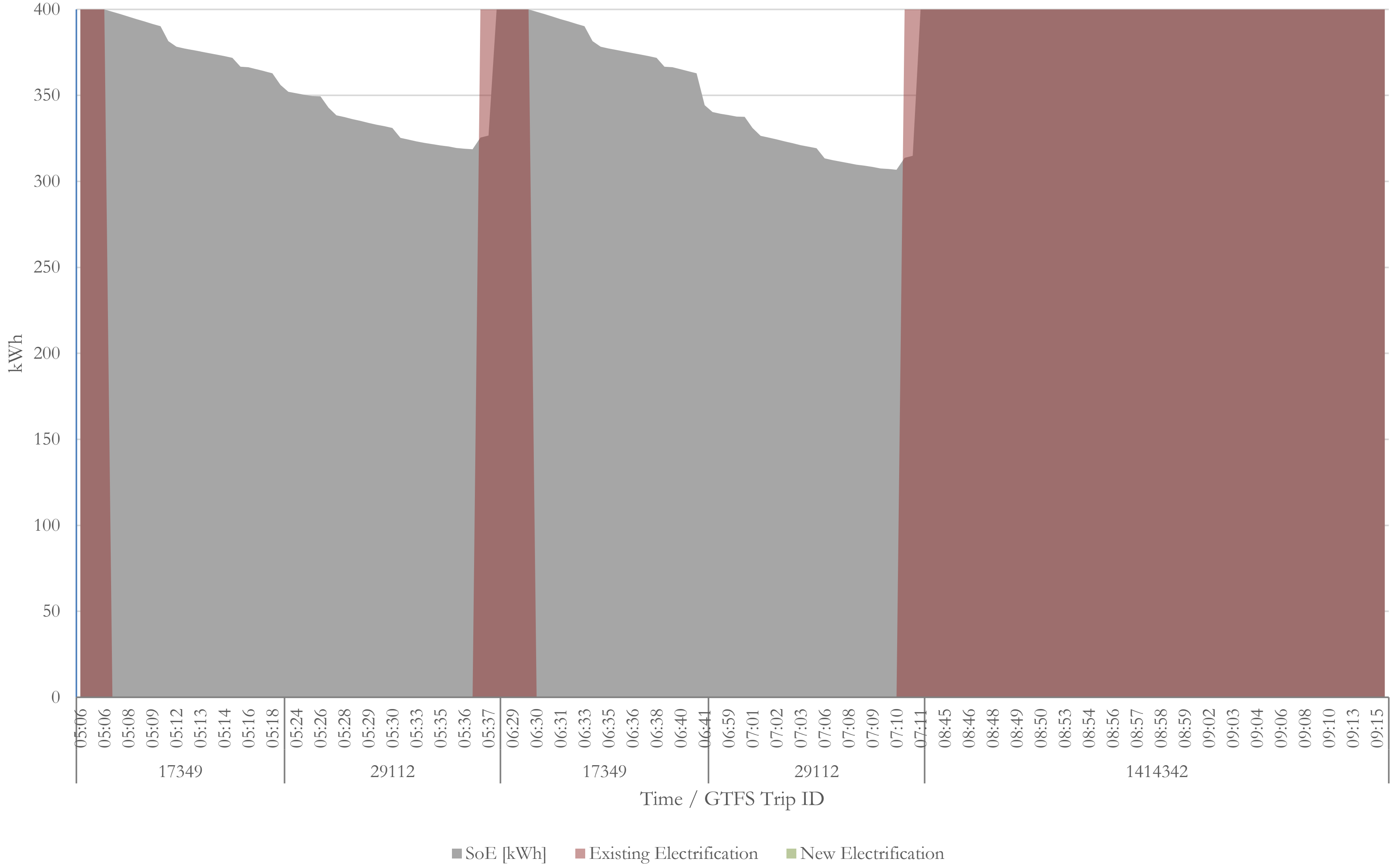

Figure 24: State of energy and the presence of electrification for vehicle 156632709 through its circulation given the distribution of existing and new electrification within the results of model run "entire Pfalznetz.

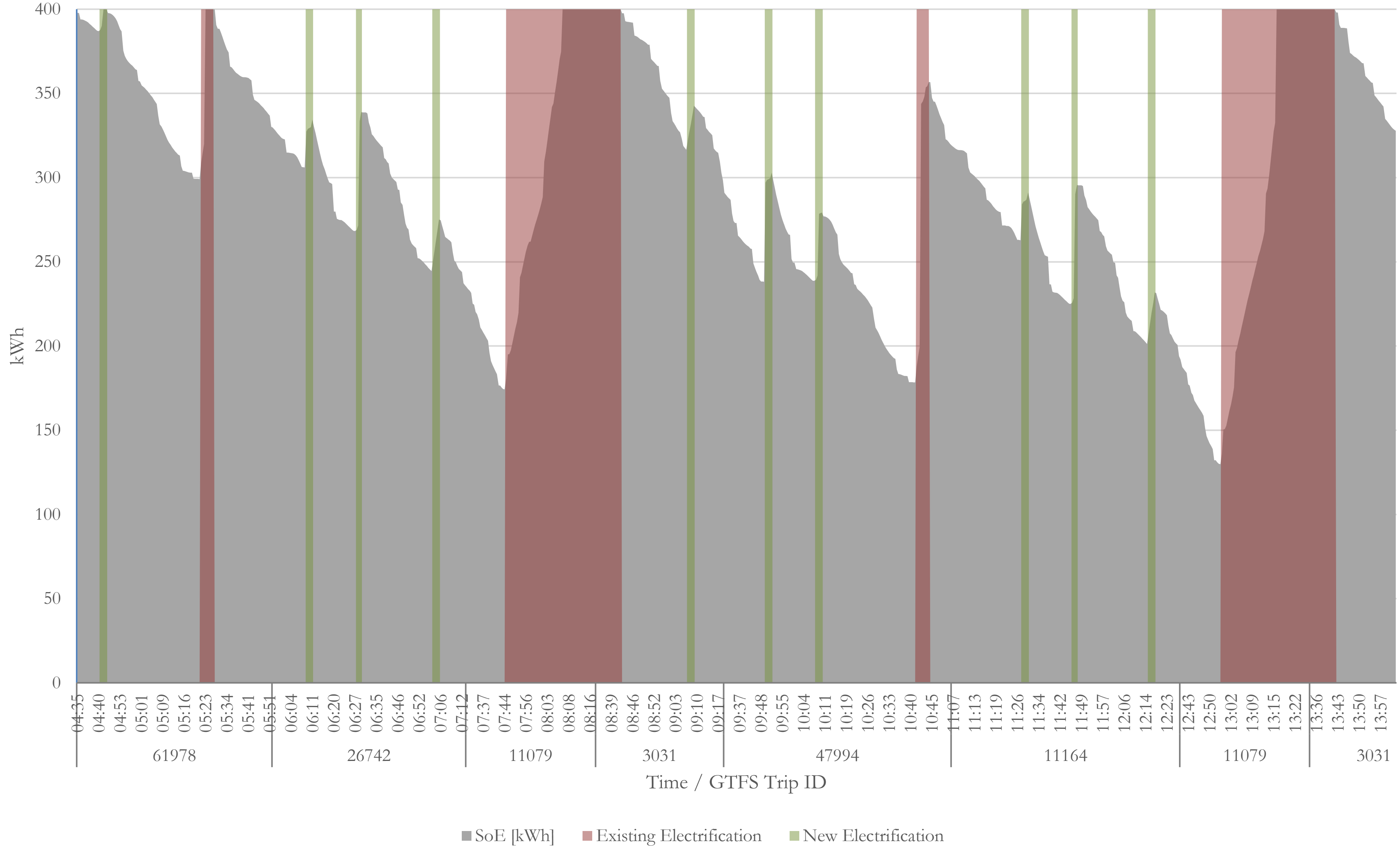

Figure 25: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

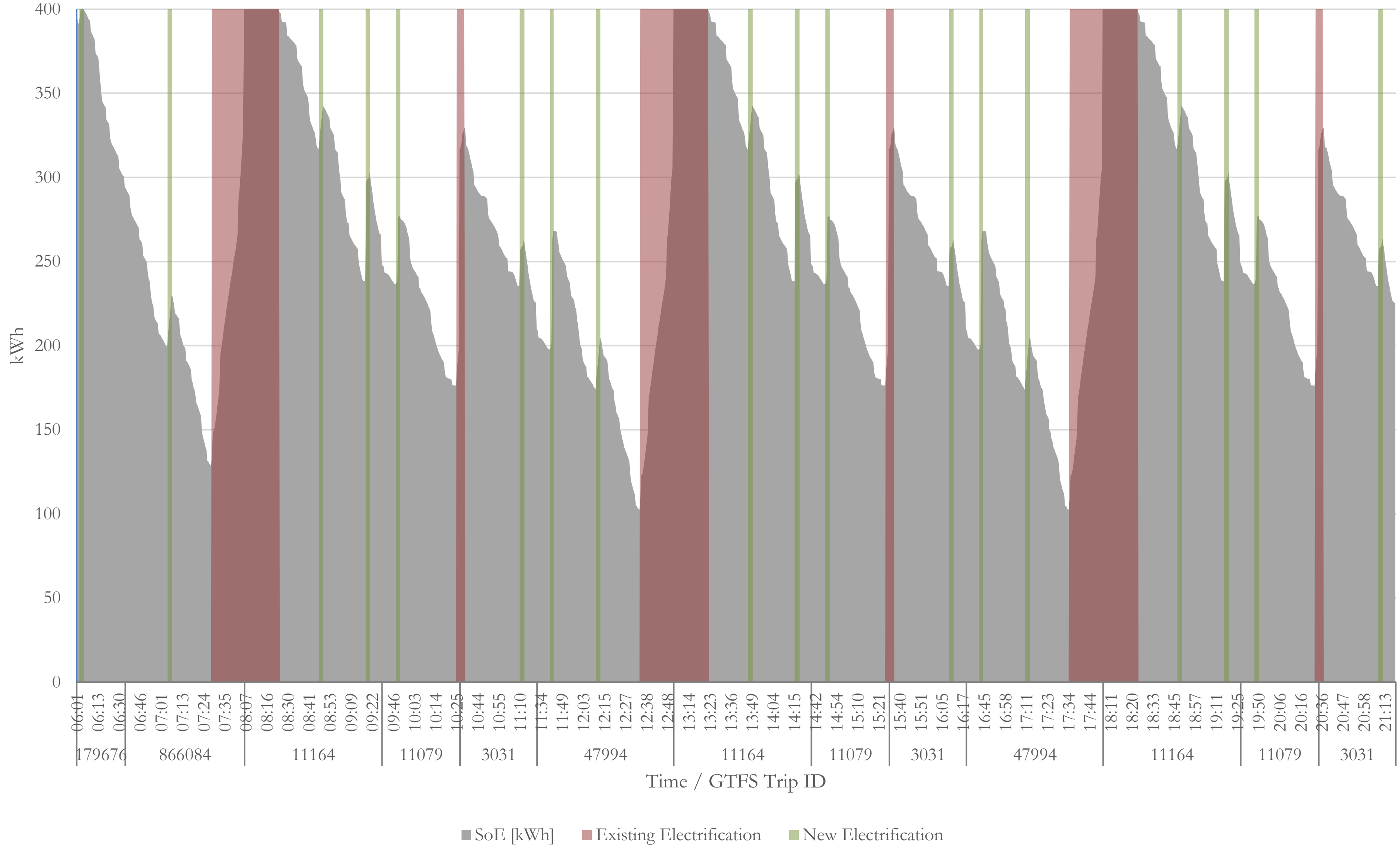

Figure 26: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

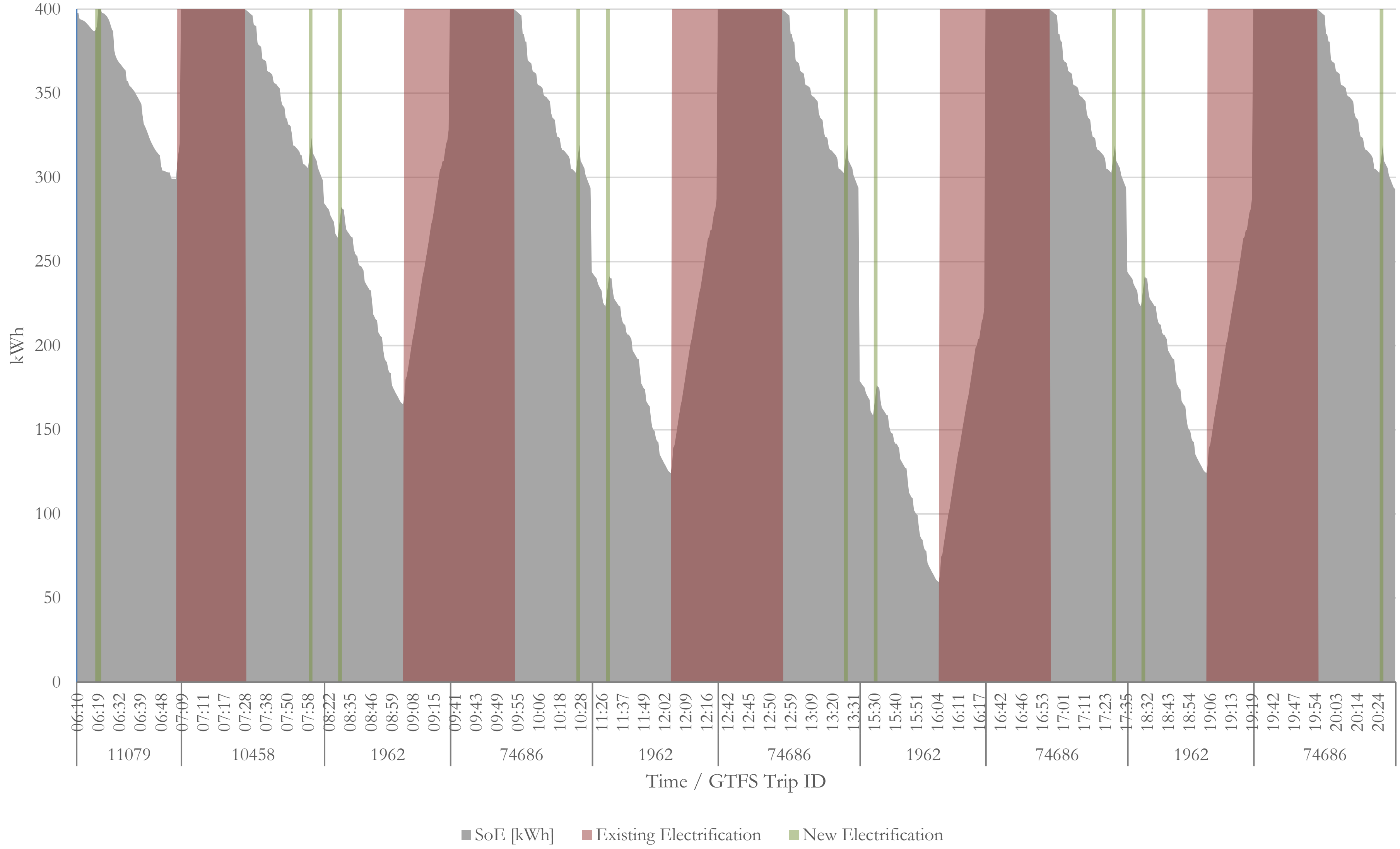

Figure 27: State of energy and the presence of electrification for vehicle 156632712 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

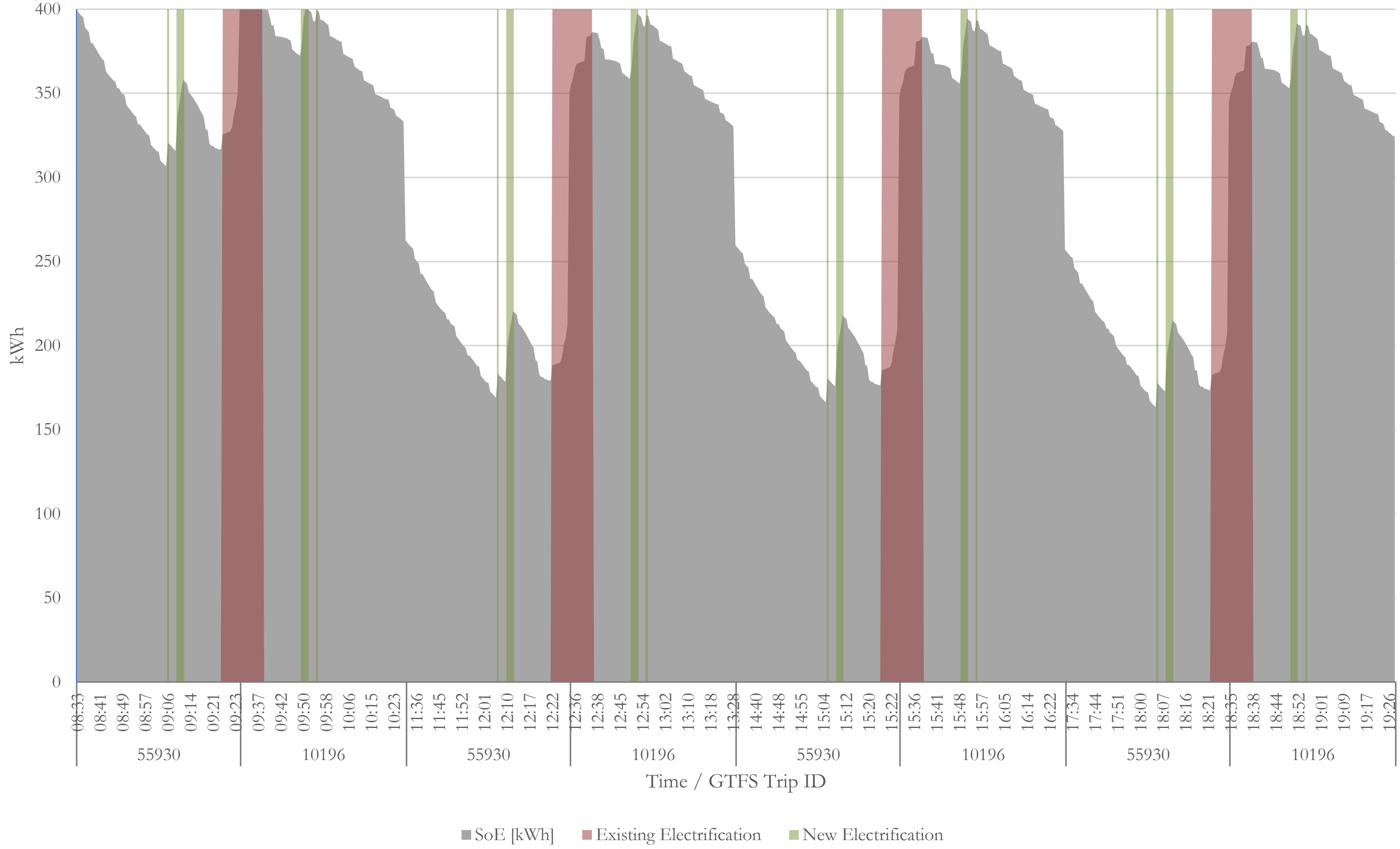

Figure 28: State of energy and the presence of electrification for vehicle 156632713 through its circulation given the distribution of existing and new electrification within the results of model run "entire Pfalznetz.

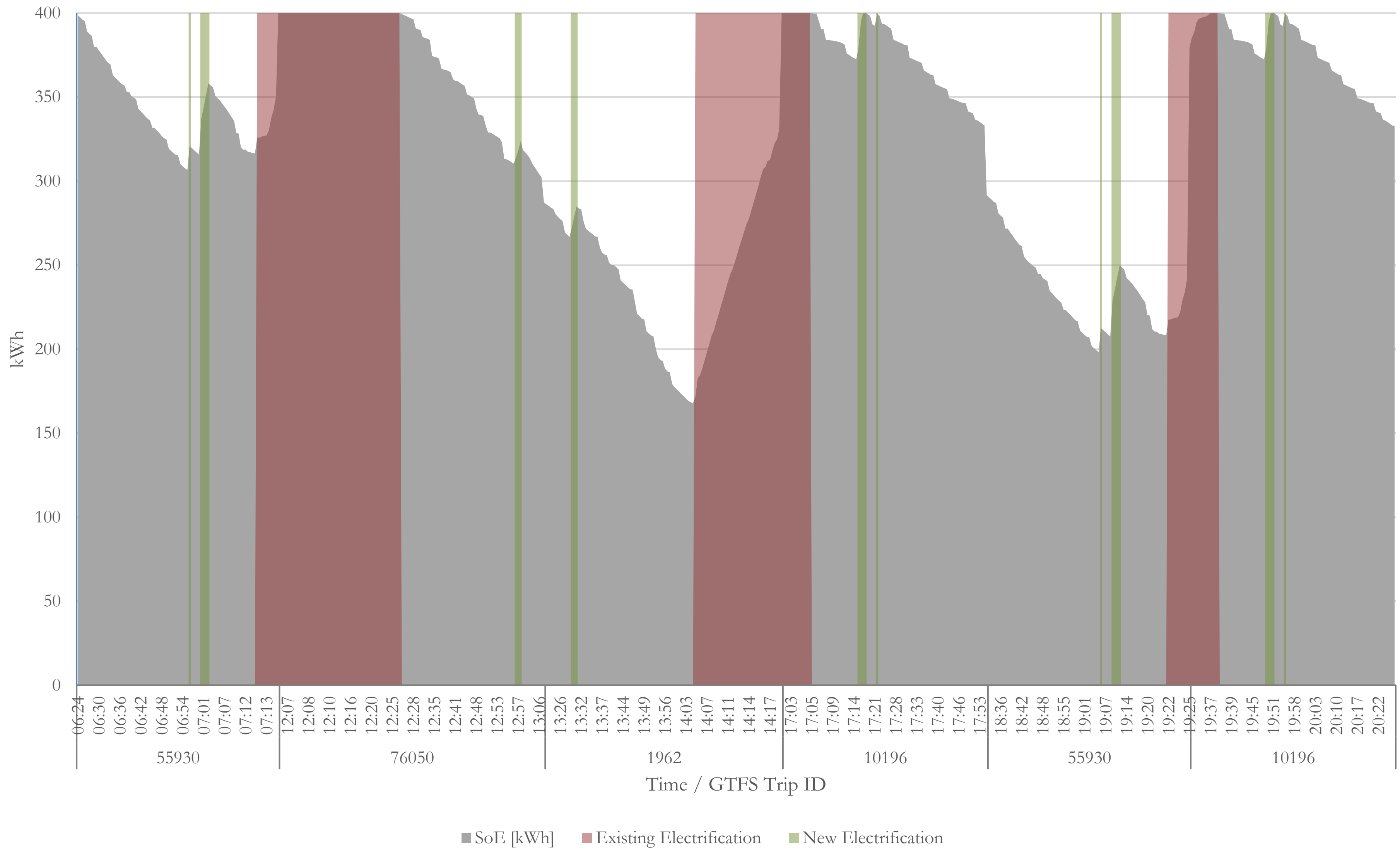

Figure 29: State of energy and the presence of electrification for vehicle 156632714 through its circulation given the distribution of existing and new electrification within the results of model run "entire Pfalznetz.

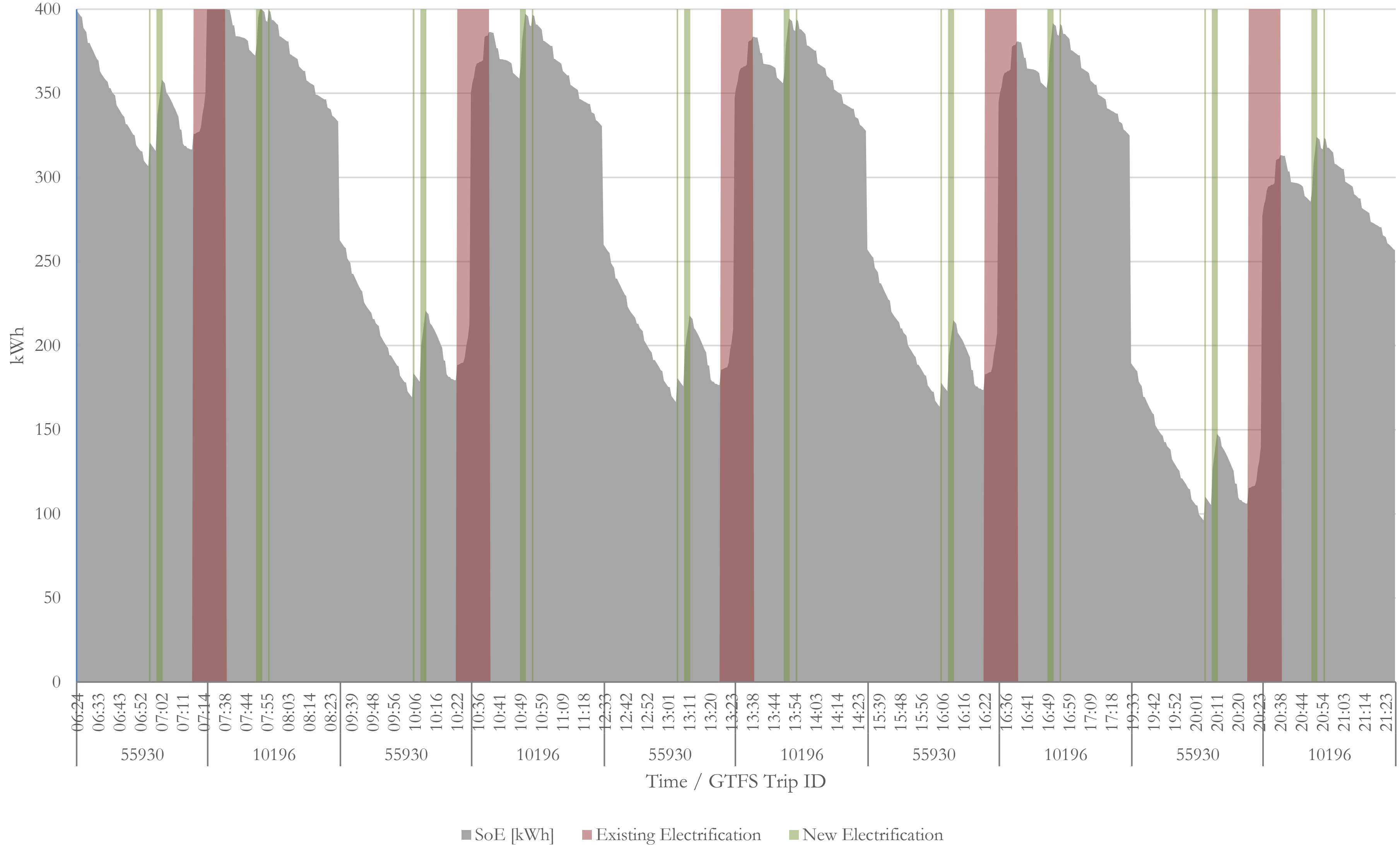

Figure 30: State of energy and the presence of electrification for vehicle 156632715 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

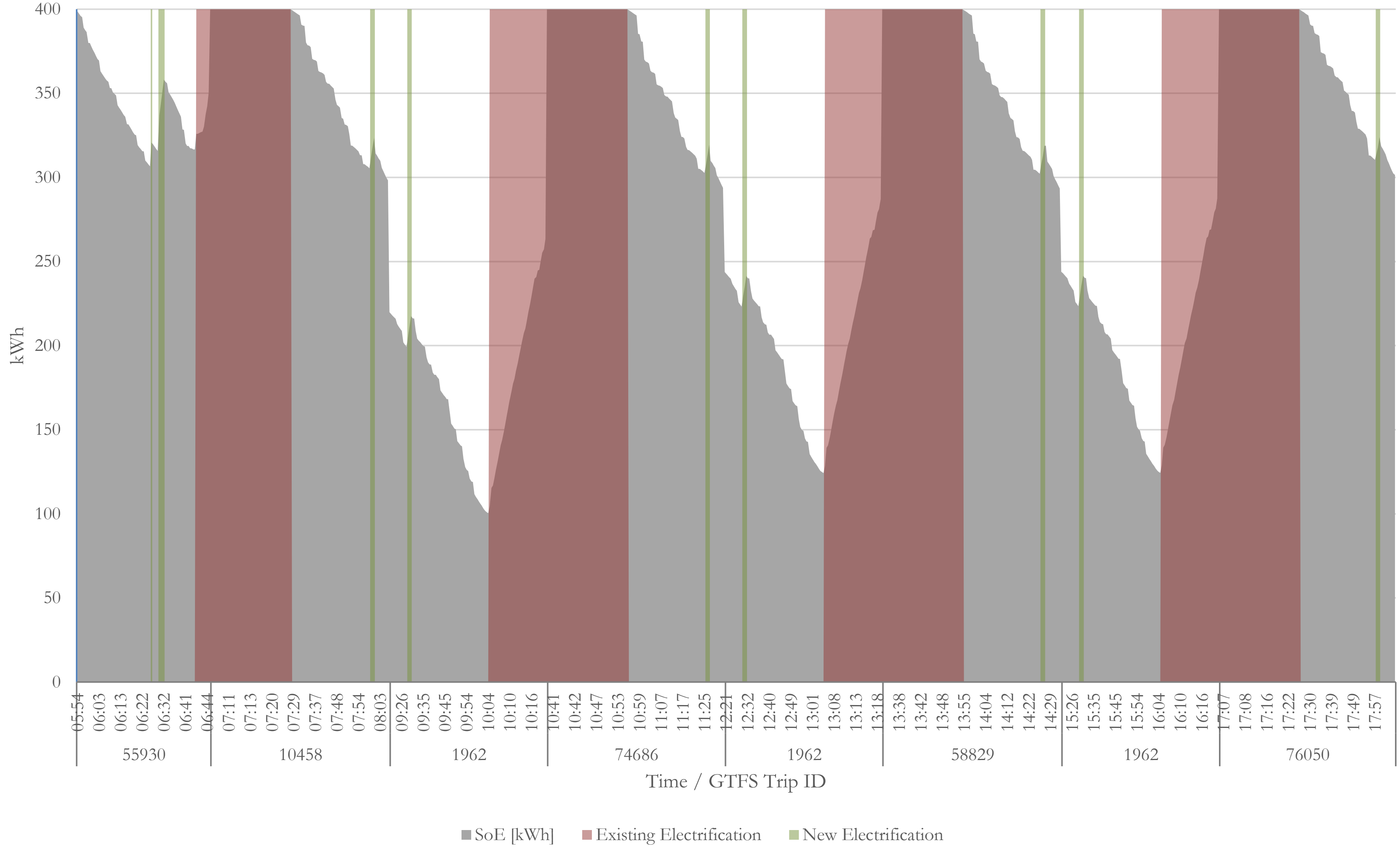

Figure 31: State of energy and the presence of electrification for vehicle 156632716 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

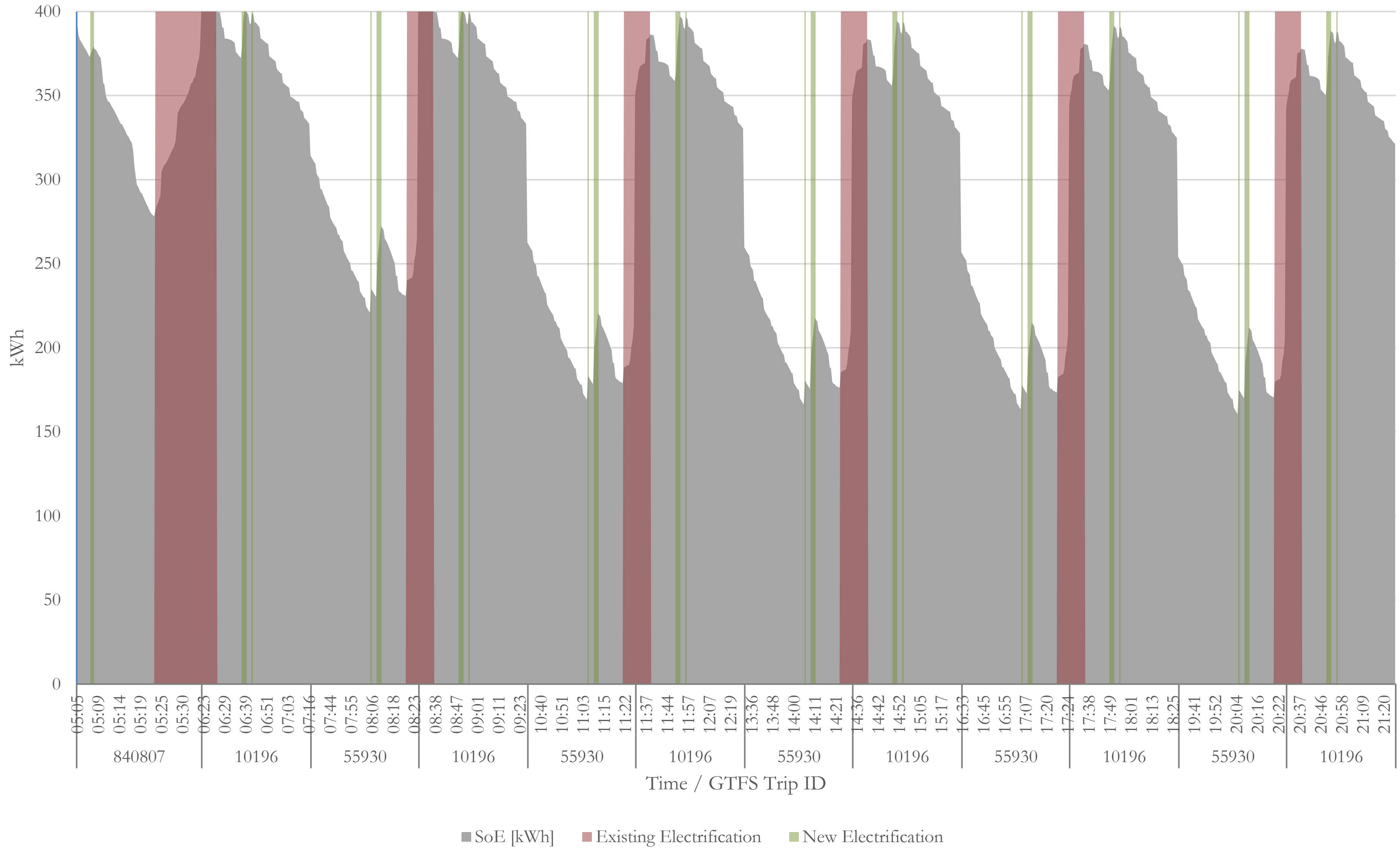

Figure 32: State of energy and the presence of electrification for vehicle 156632717 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

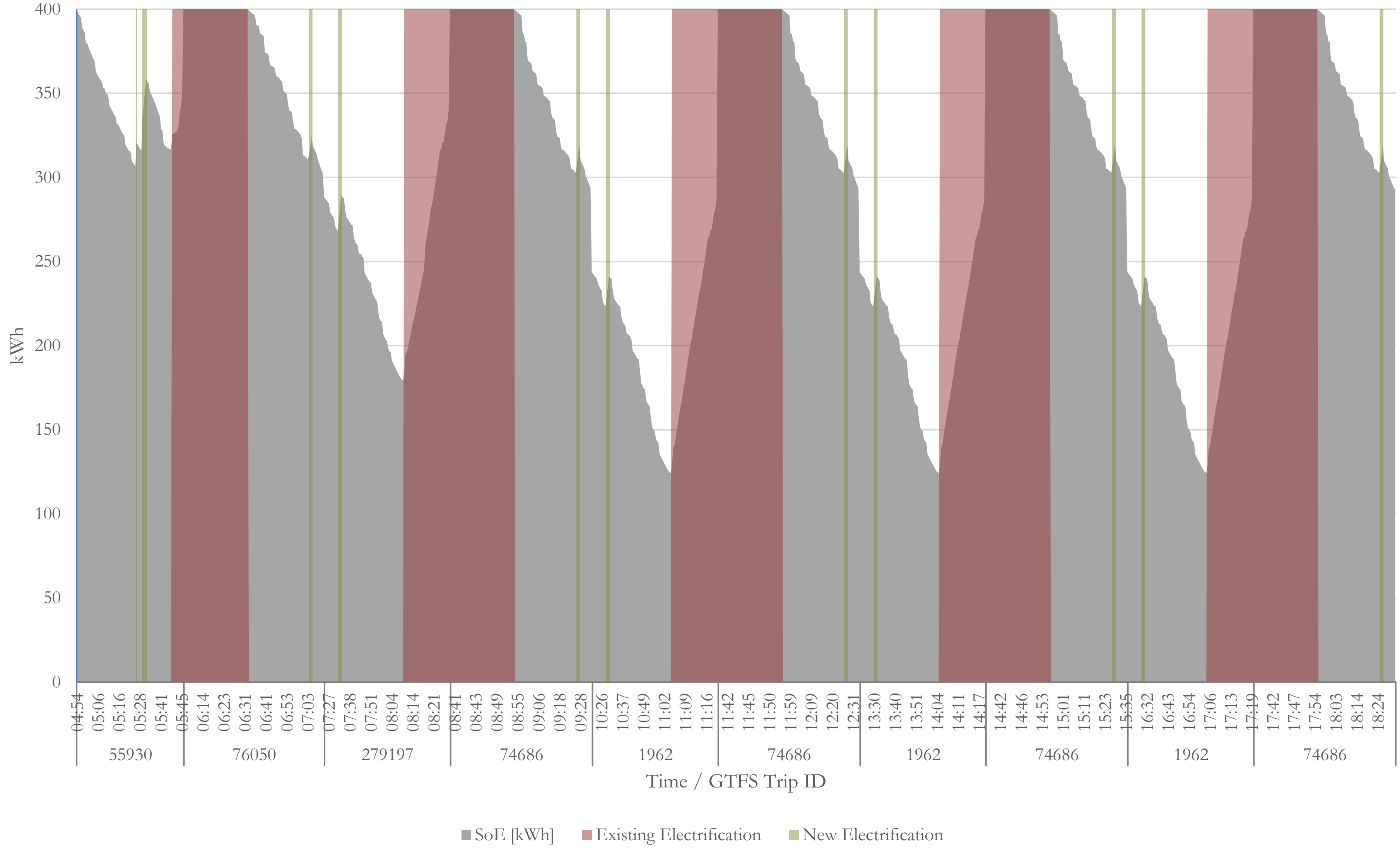

Figure 33: State of energy and the presence of electrification for vehicle 156632718 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

Figure 34: State of energy and the presence of electrification for vehicle 156632719 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

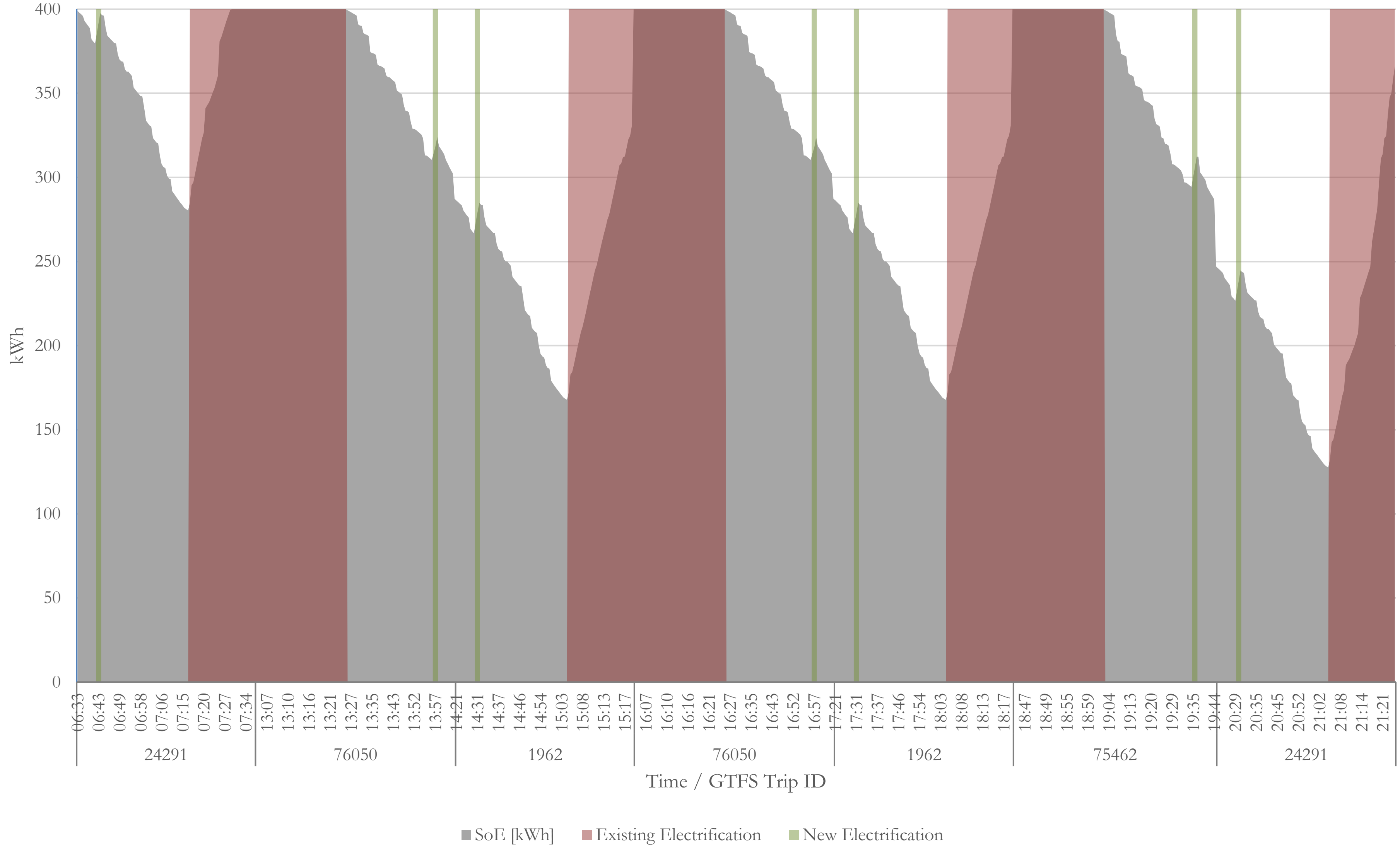

Figure 35: State of energy and the presence of electrification for vehicle 156632720 through its circulation given the distribution of existing and new electrification within the results of model run "entire Pfalznetz.

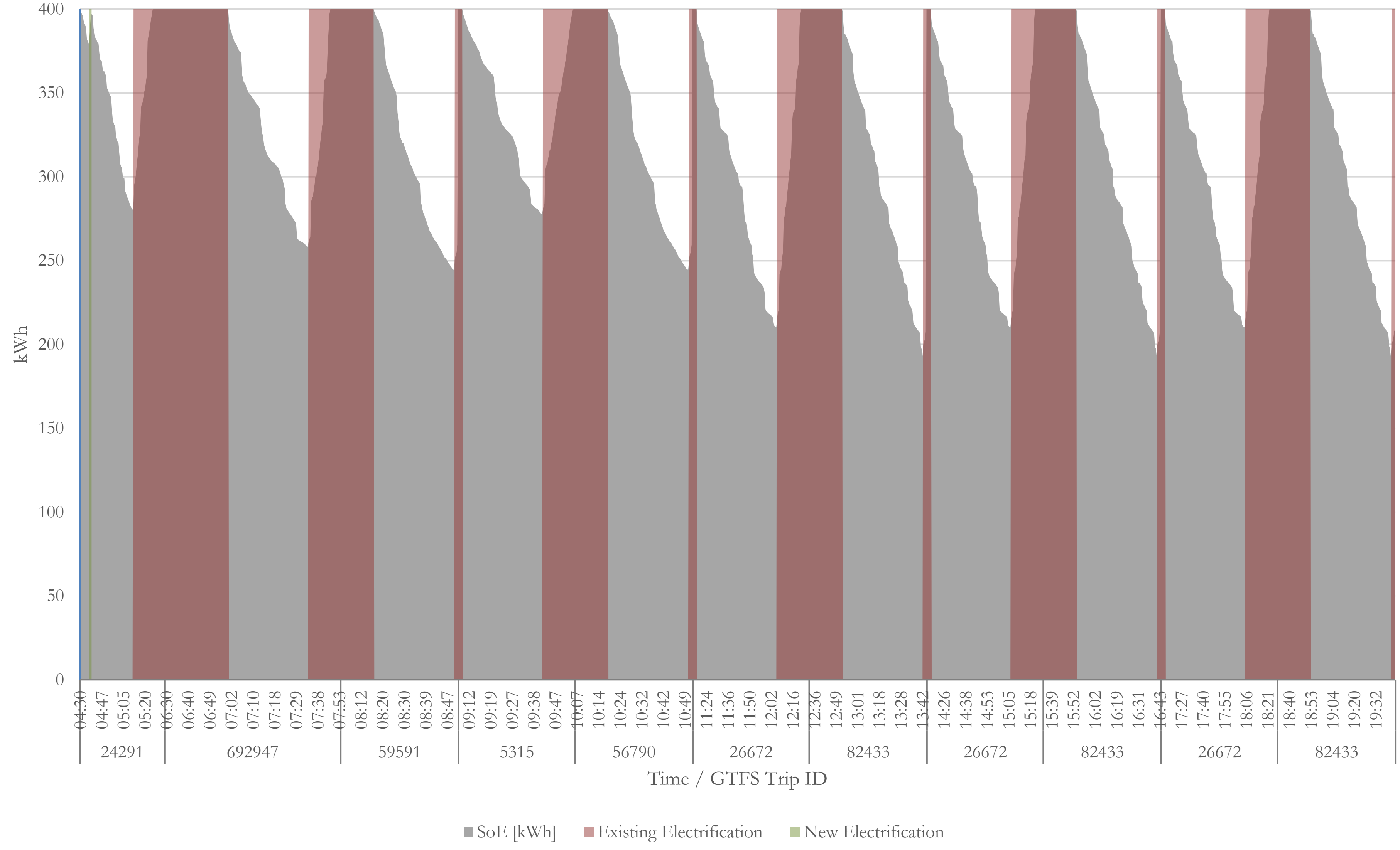

Figure 36: State of energy and the presence of electrification for vehicle 156632721 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

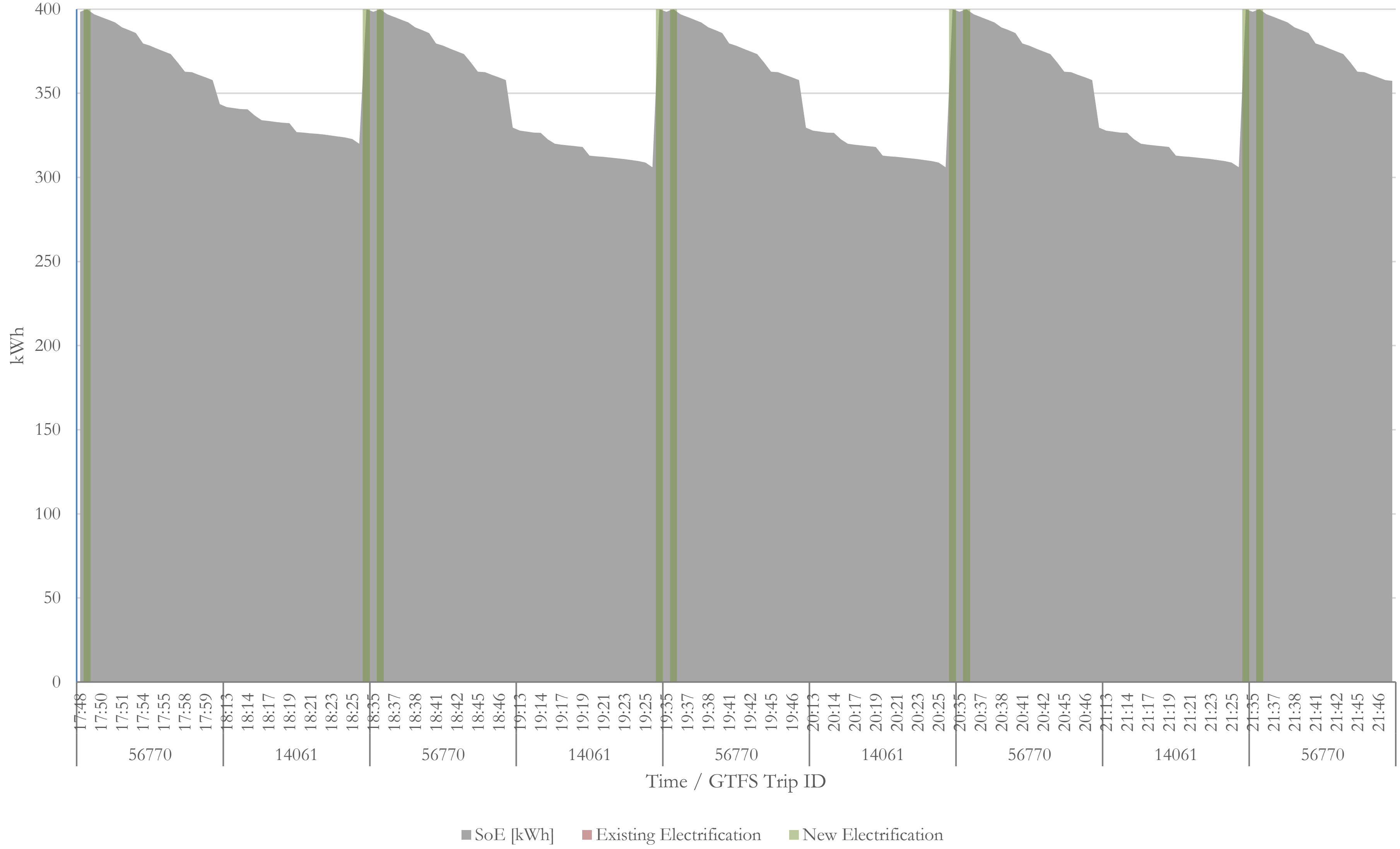

Figure 37: State of energy and the presence of electrification for vehicle 156632722 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

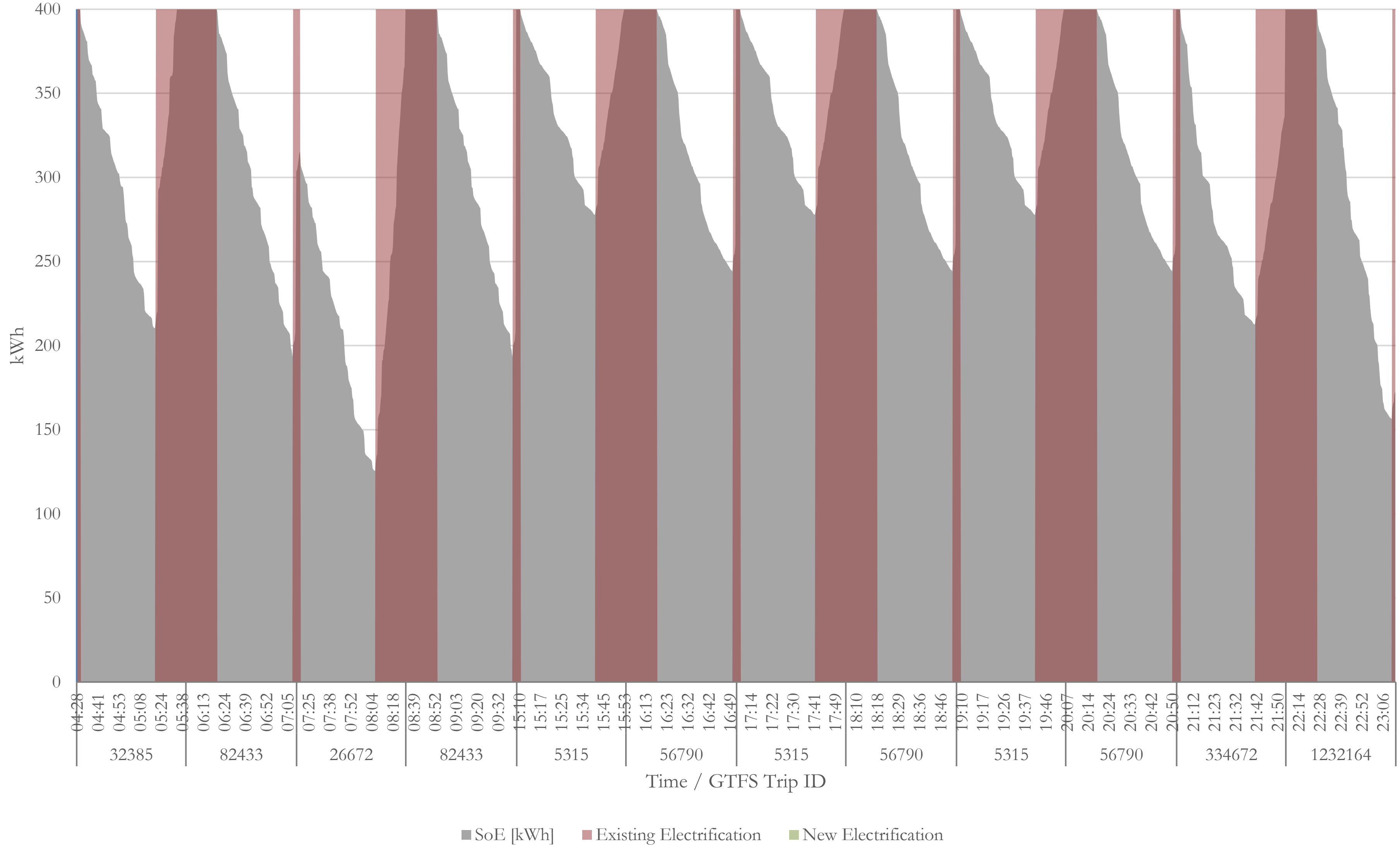

Figure 38: State of energy and the presence of electrification for vehicle 156632724 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

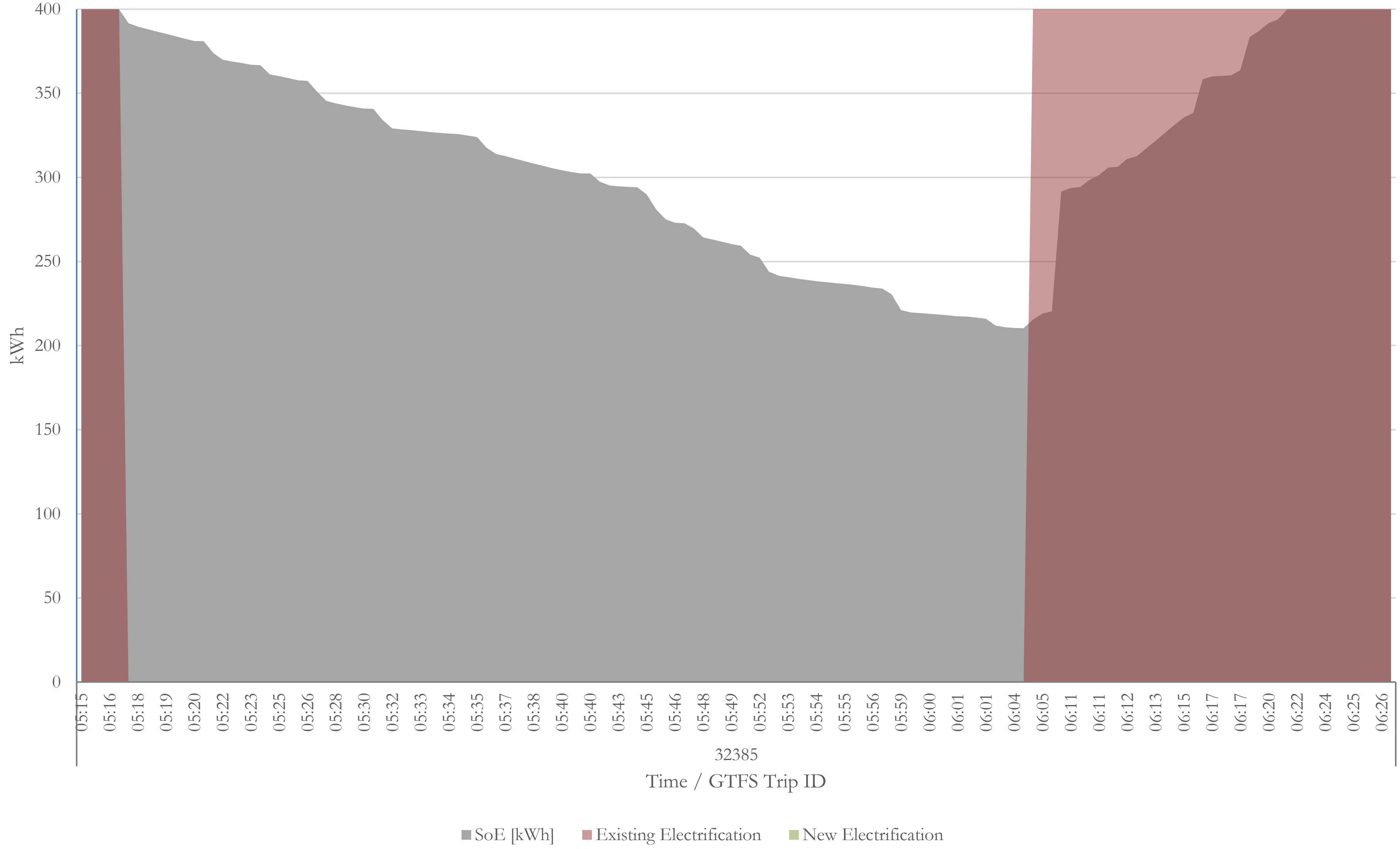

Figure 39: State of energy and the presence of electrification for vehicle 156632725 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

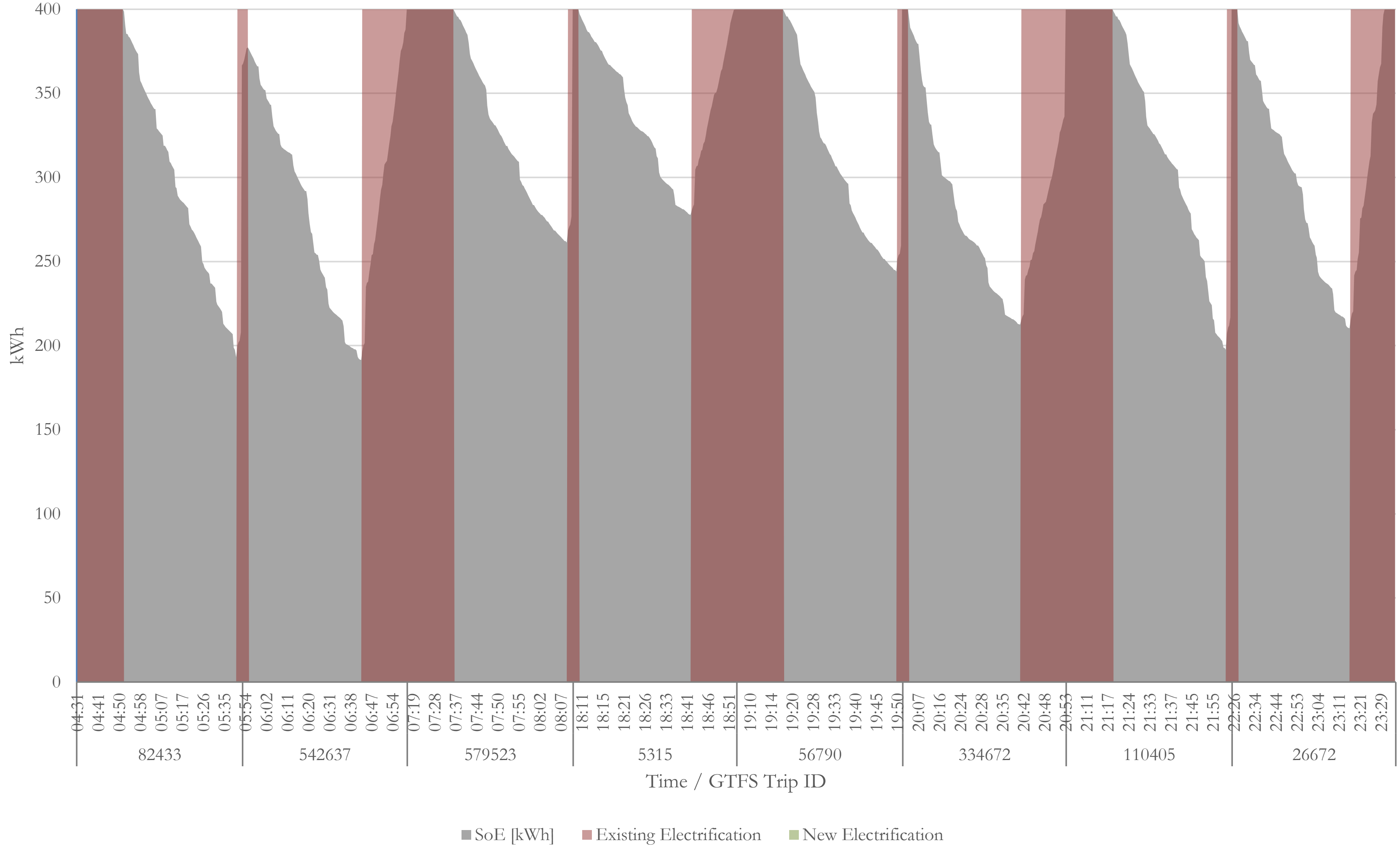

Figure 40: State of energy and the presence of electrification for vehicle 156632726 through its circulation given the distribution of existing and new electrification within the results of model run "entire Pfalznetz.

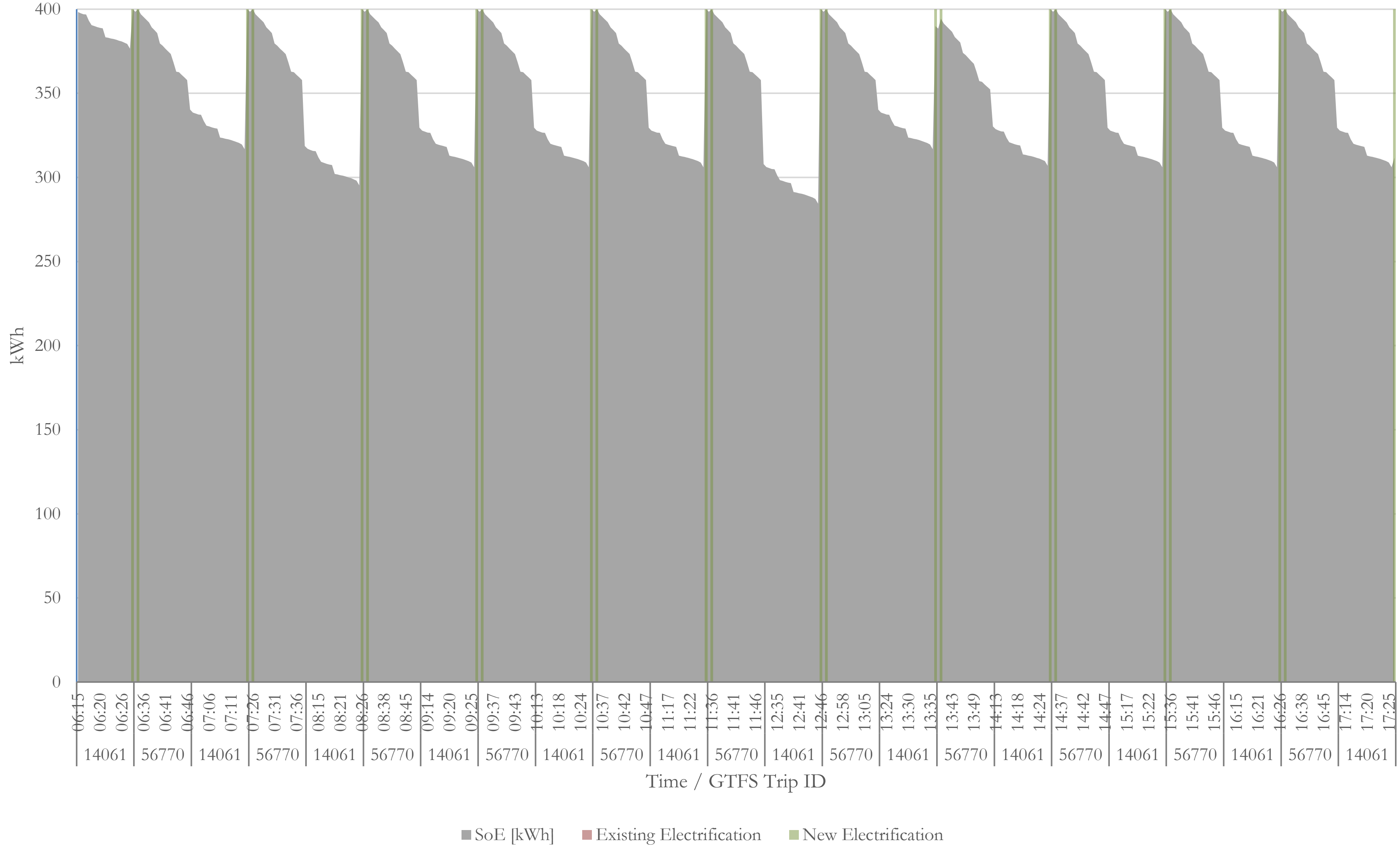

Figure 41: State of energy and the presence of electrification for vehicle 156632727 through its circulation given the distribution of existing and new electrification within the results of model run "entire Pfalznetz.

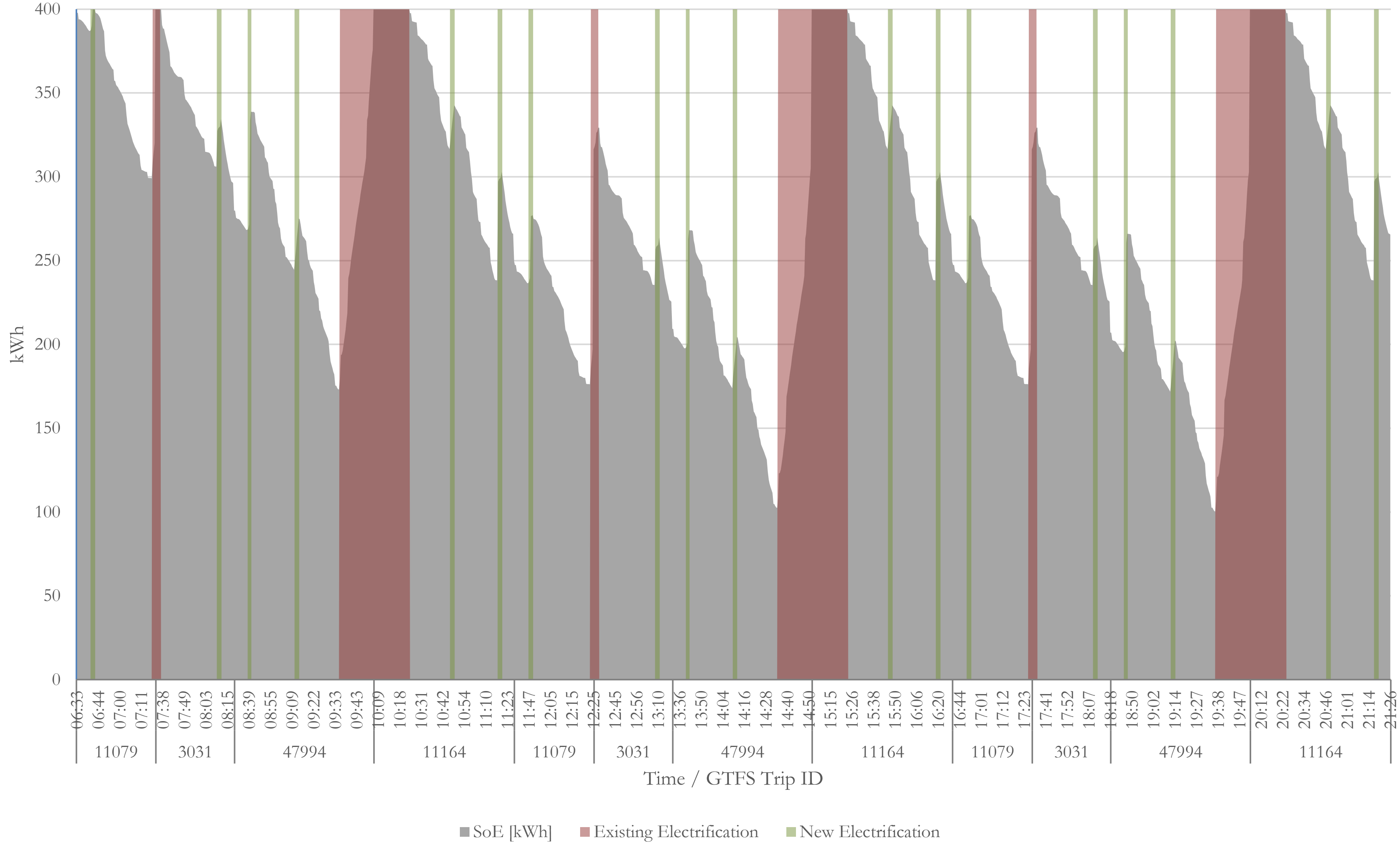

Figure 42: State of energy and the presence of electrification for vehicle 156632730 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

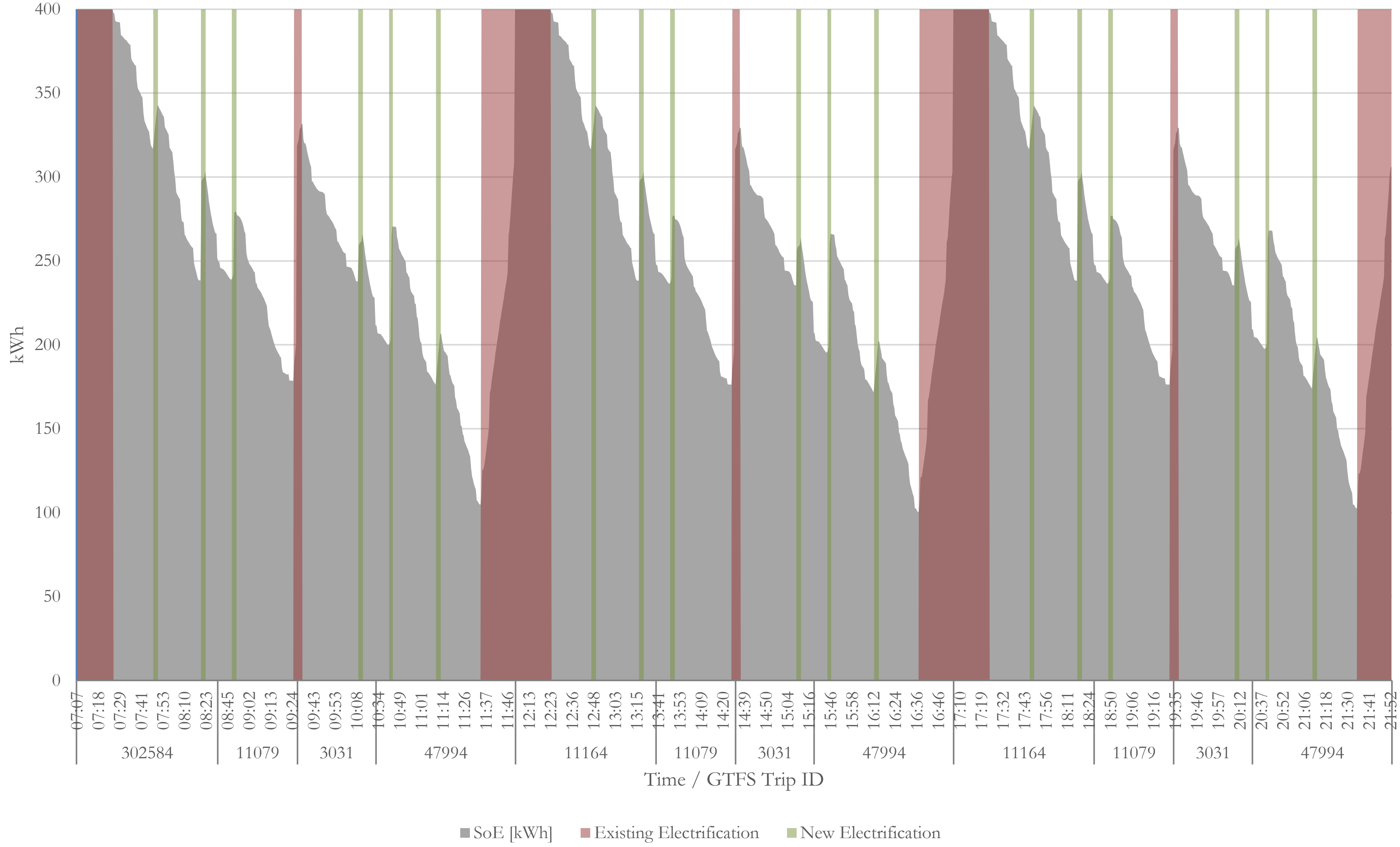

Figure 43: State of energy and the presence of electrification for vehicle 156632732 through its circulation given the distribution of existing and new electrification within the results of model run "entire Pfalznetz.

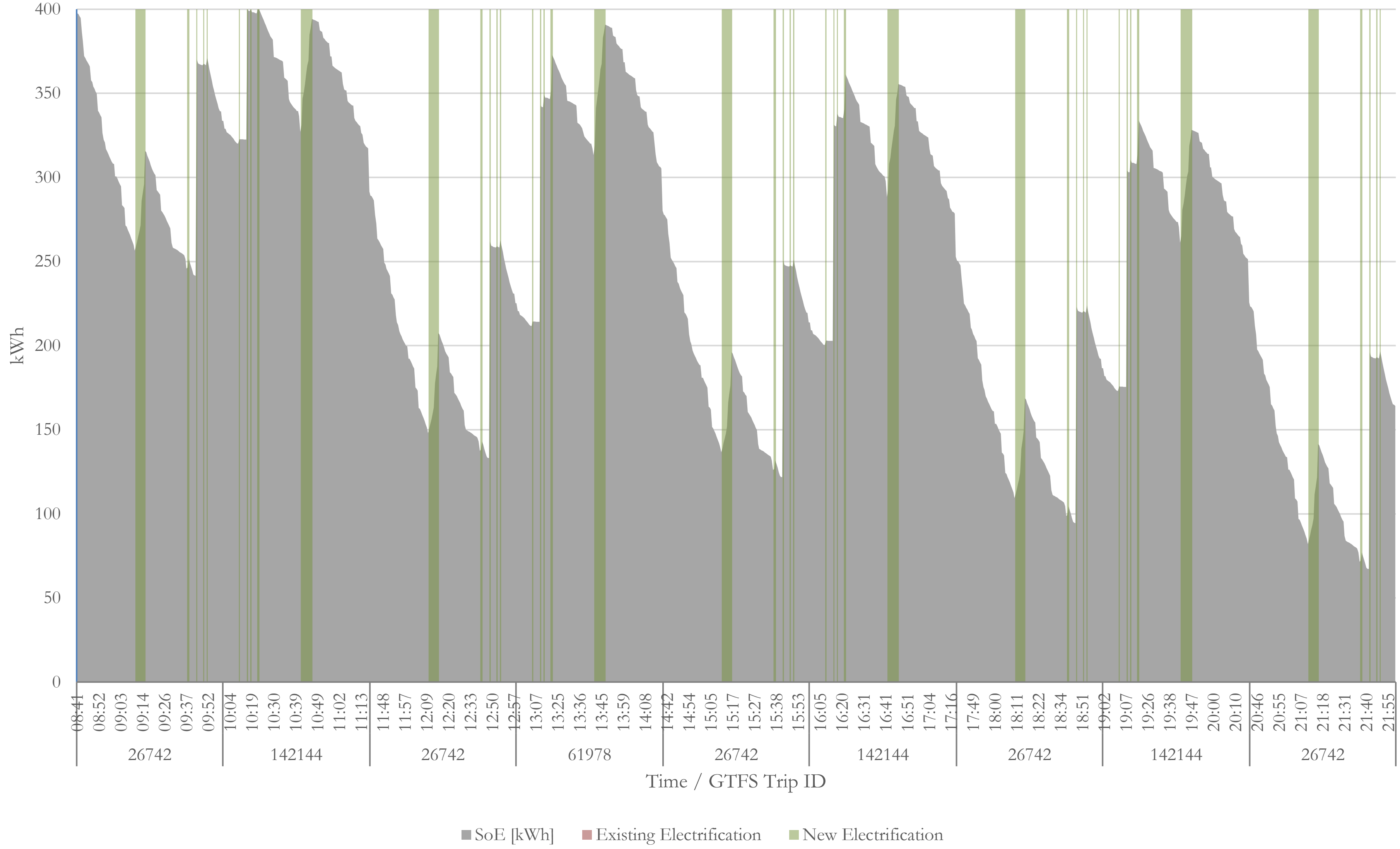

Figure 44: State of energy and the presence of electrification for vehicle 156632733 through its circulation given the distribution of existing and new electrification within the results of model run "entire Pfalznetz.

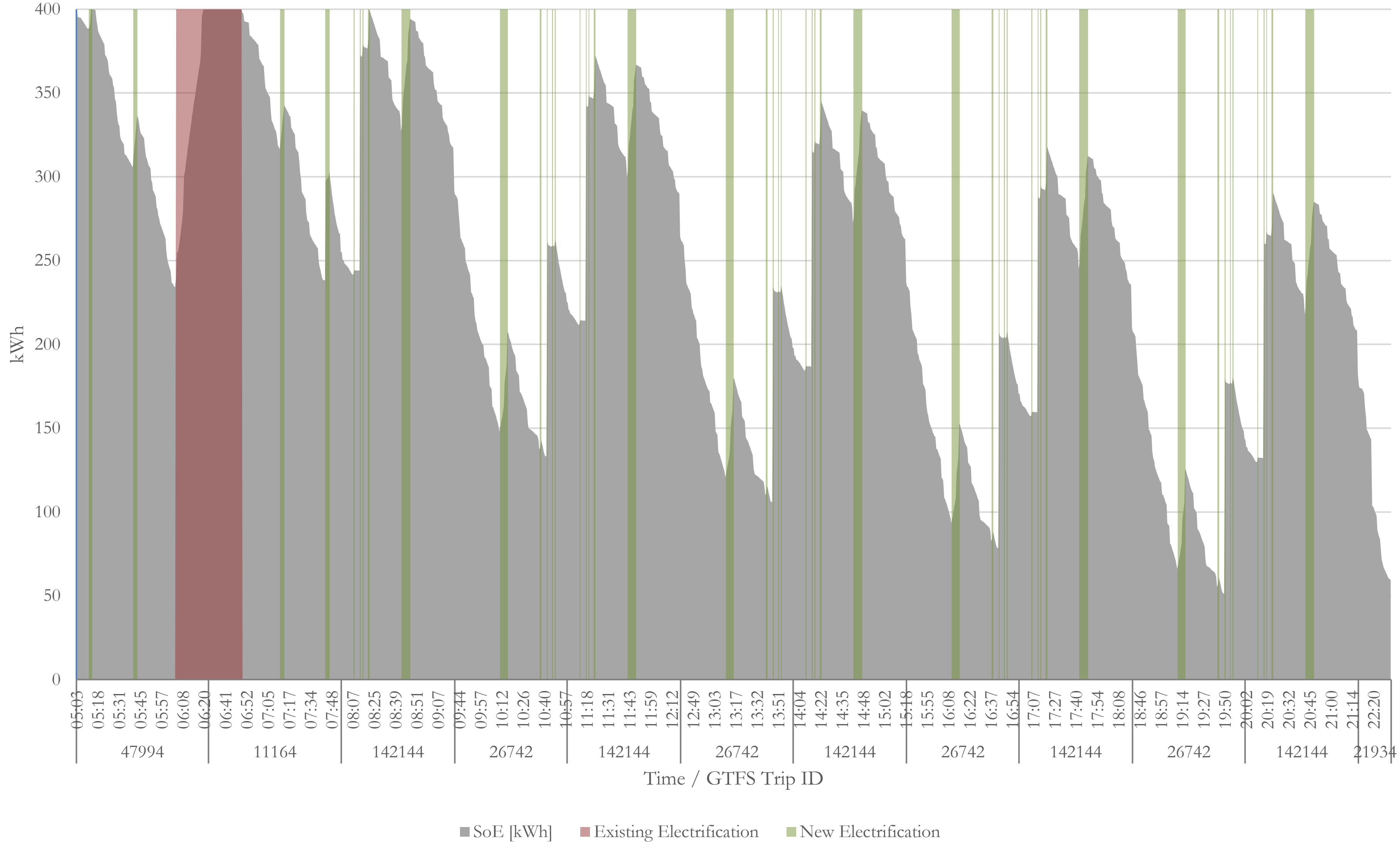

Figure 45: State of energy and the presence of electrification for vehicle 156632734 through its circulation given the distribution of existing and new electrification within the results of model run “entire Pfalznetz.

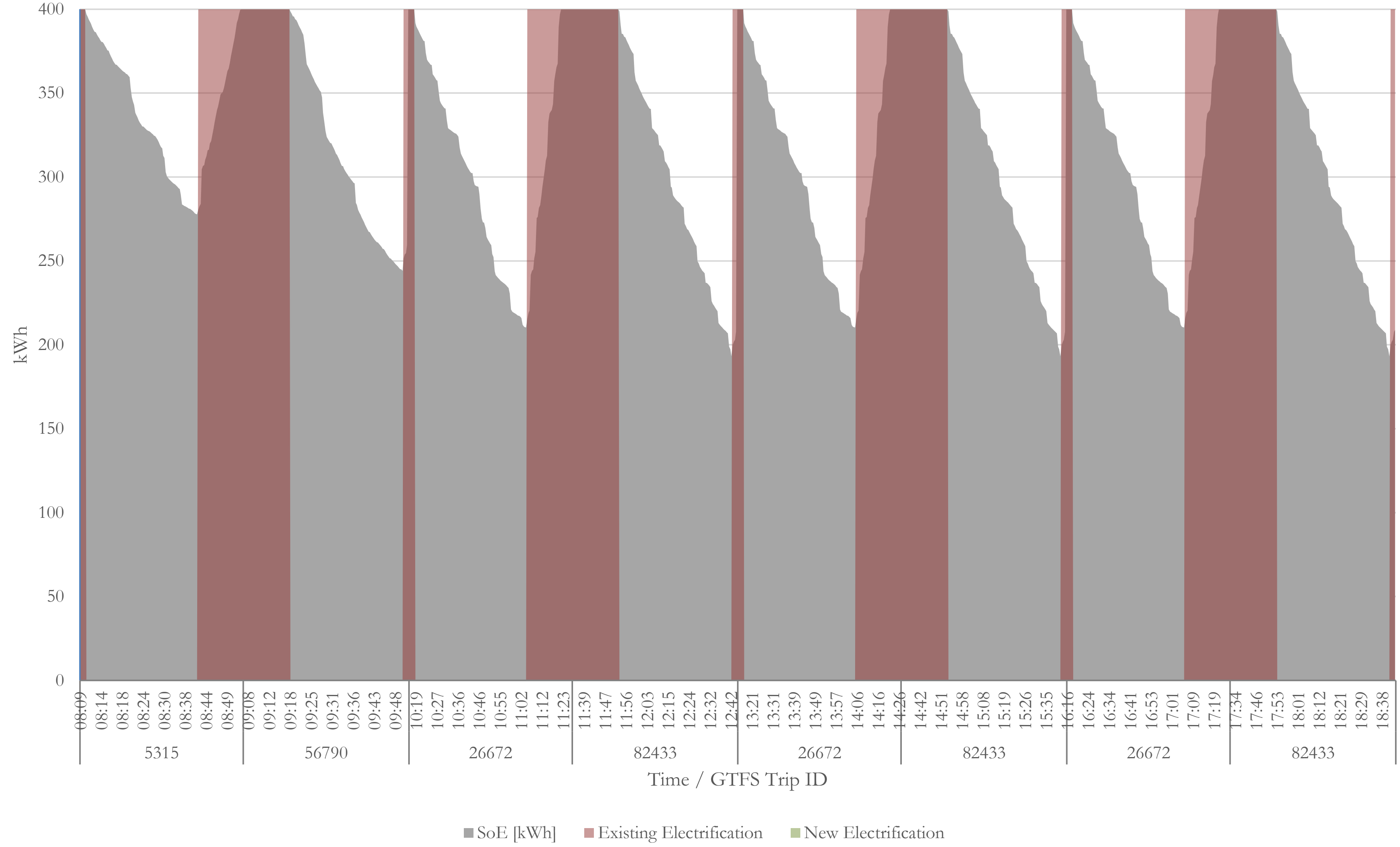

Figure 46: State of energy and the presence of electrification for vehicle 156632737 through its circulation given the distribution of existing and new electrification within the results of model run "entire Pfalznetz.

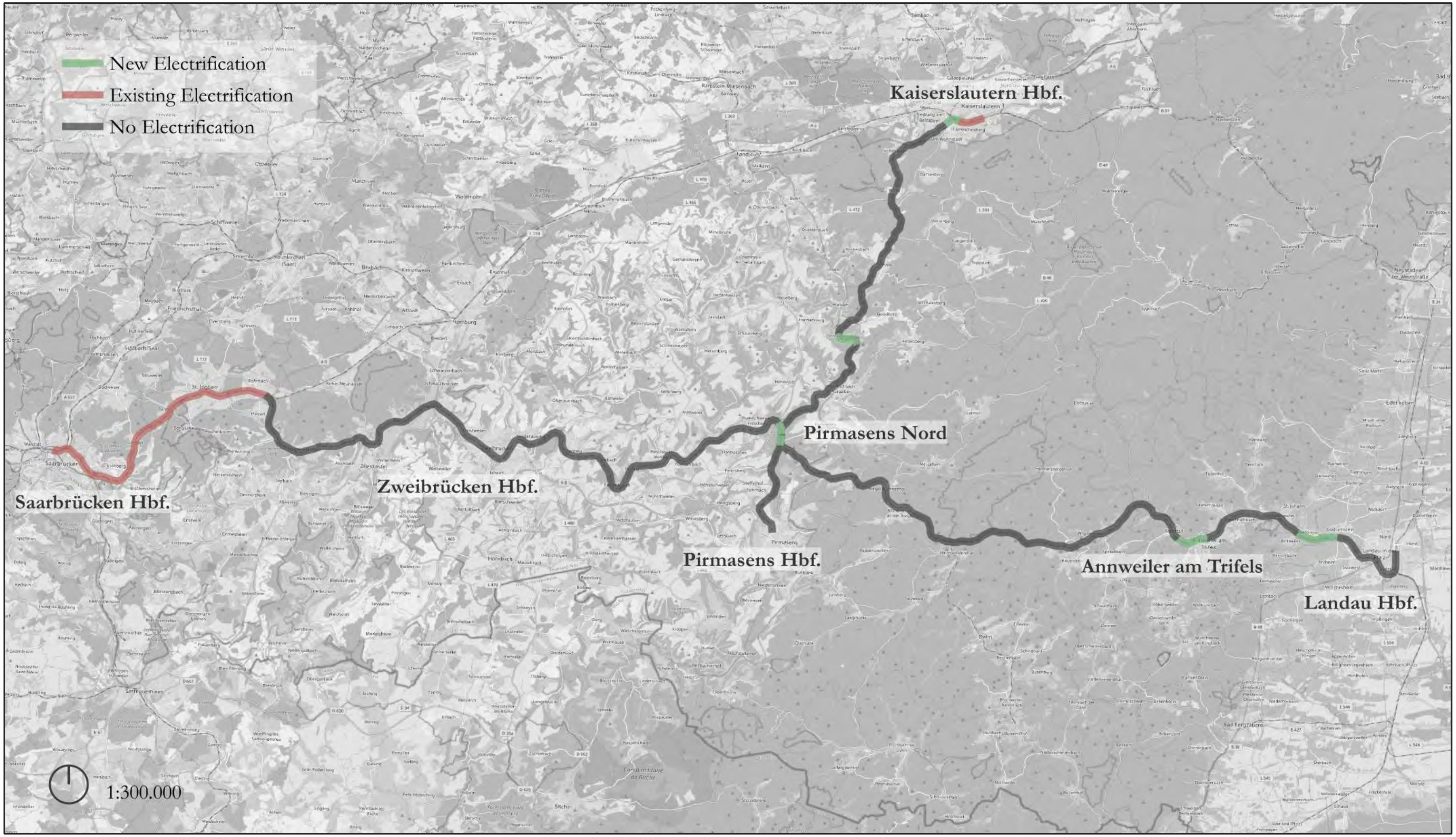

Figure 47: Existing and proposed electrification for model run “small battery 1”. The electrification is given per track even though this is not visible at map scale.

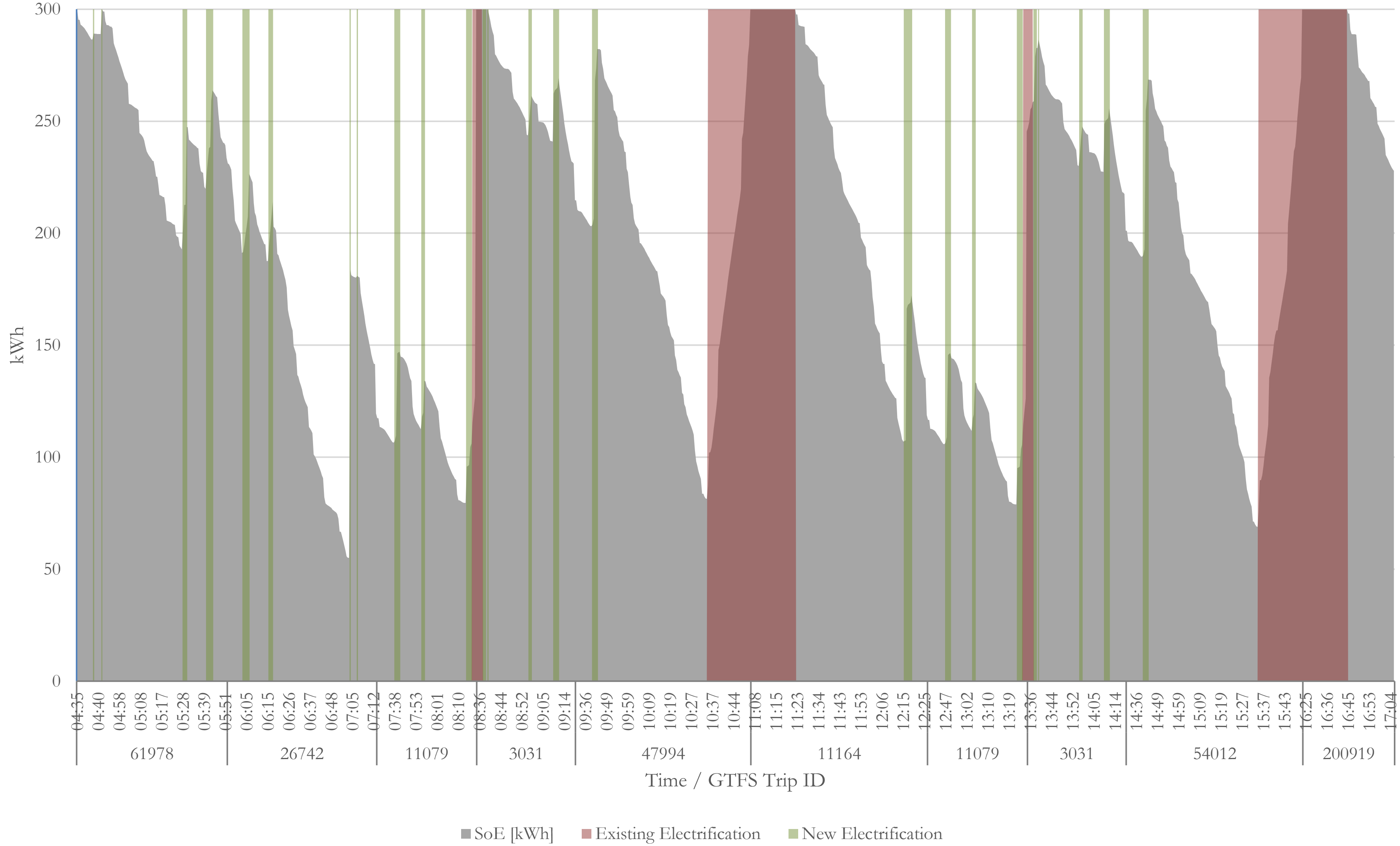

Figure 48: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run “small battery 1”.

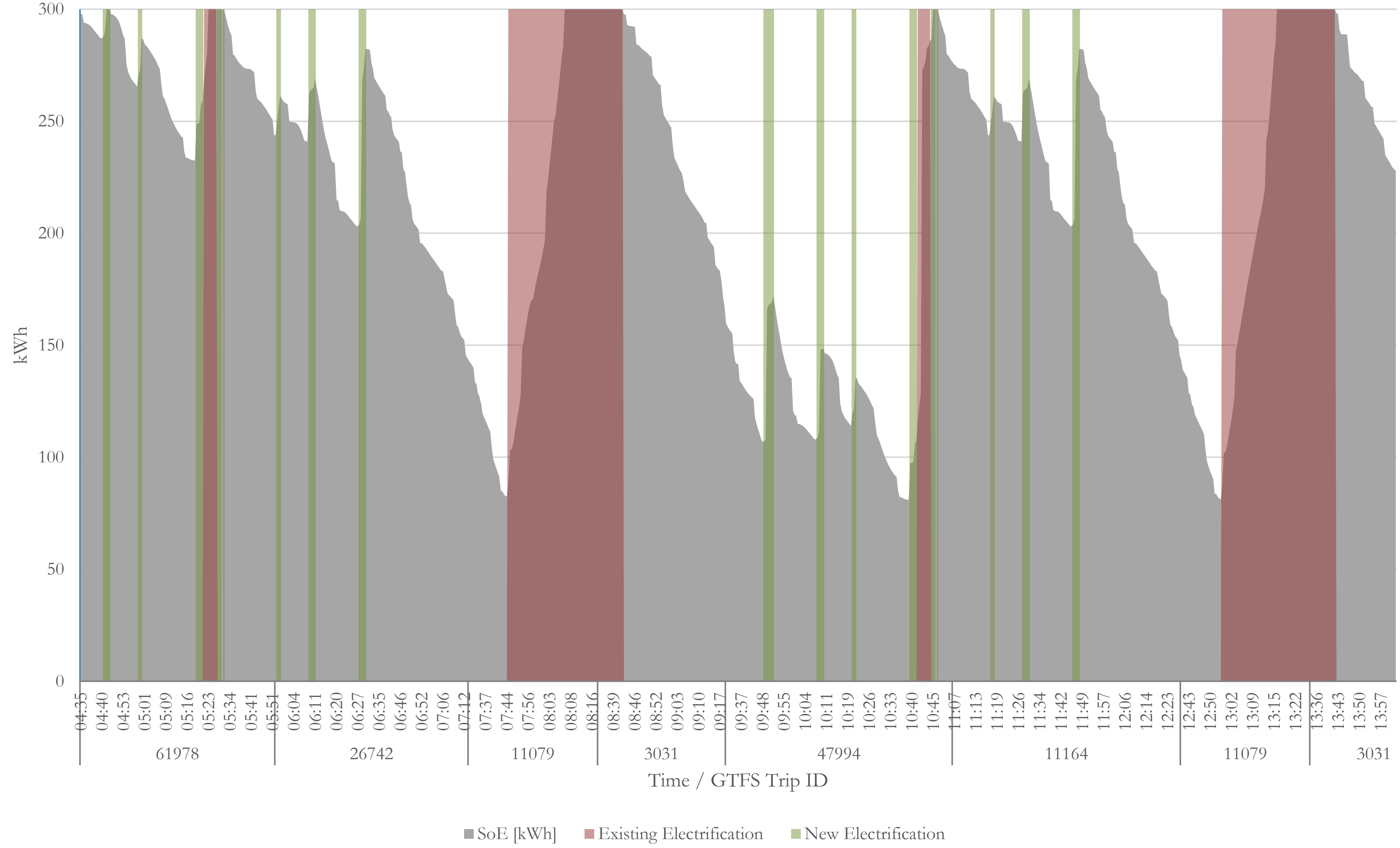

Figure 49: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run “small battery 1”.

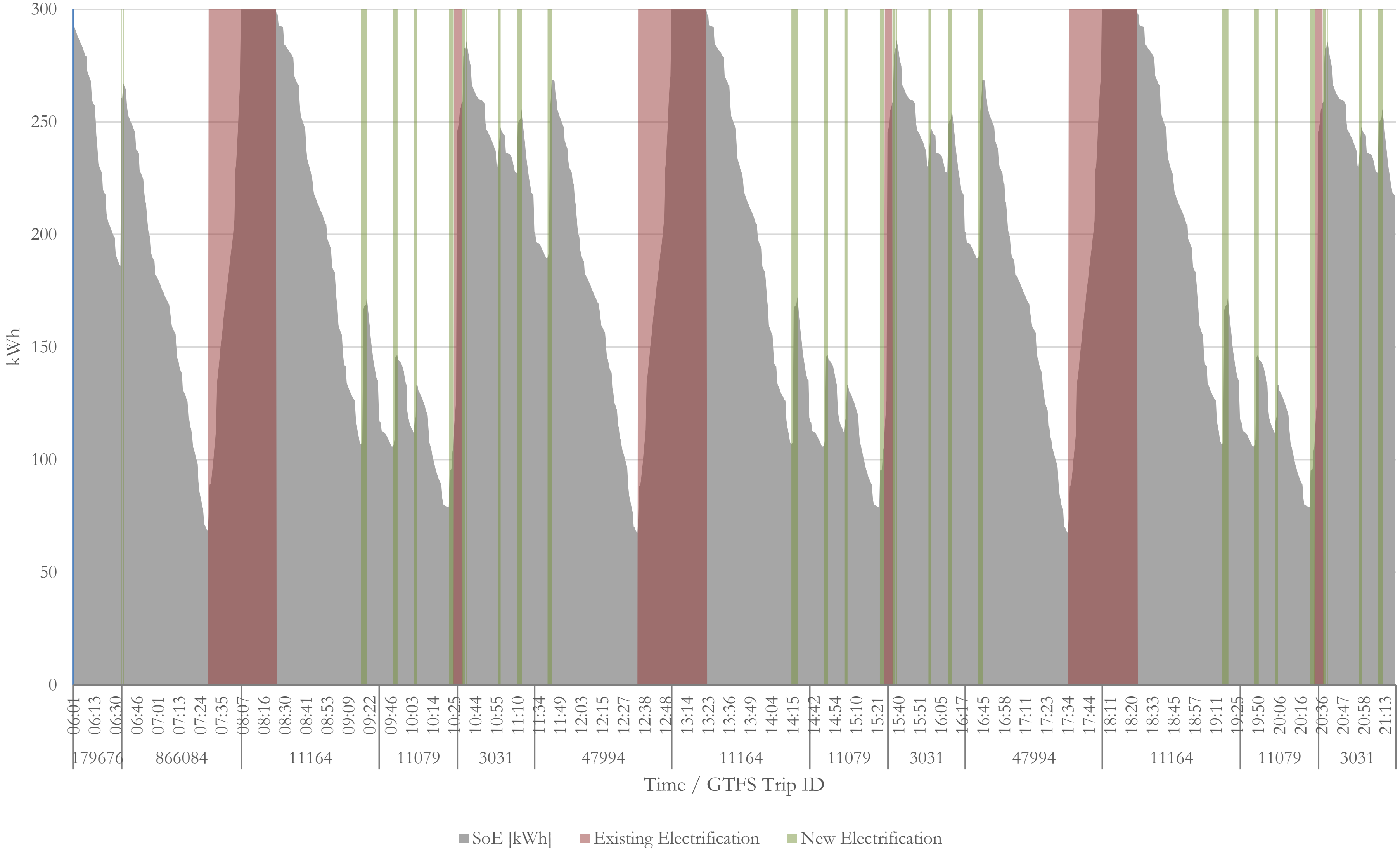

Figure 50: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run “small battery 1”.

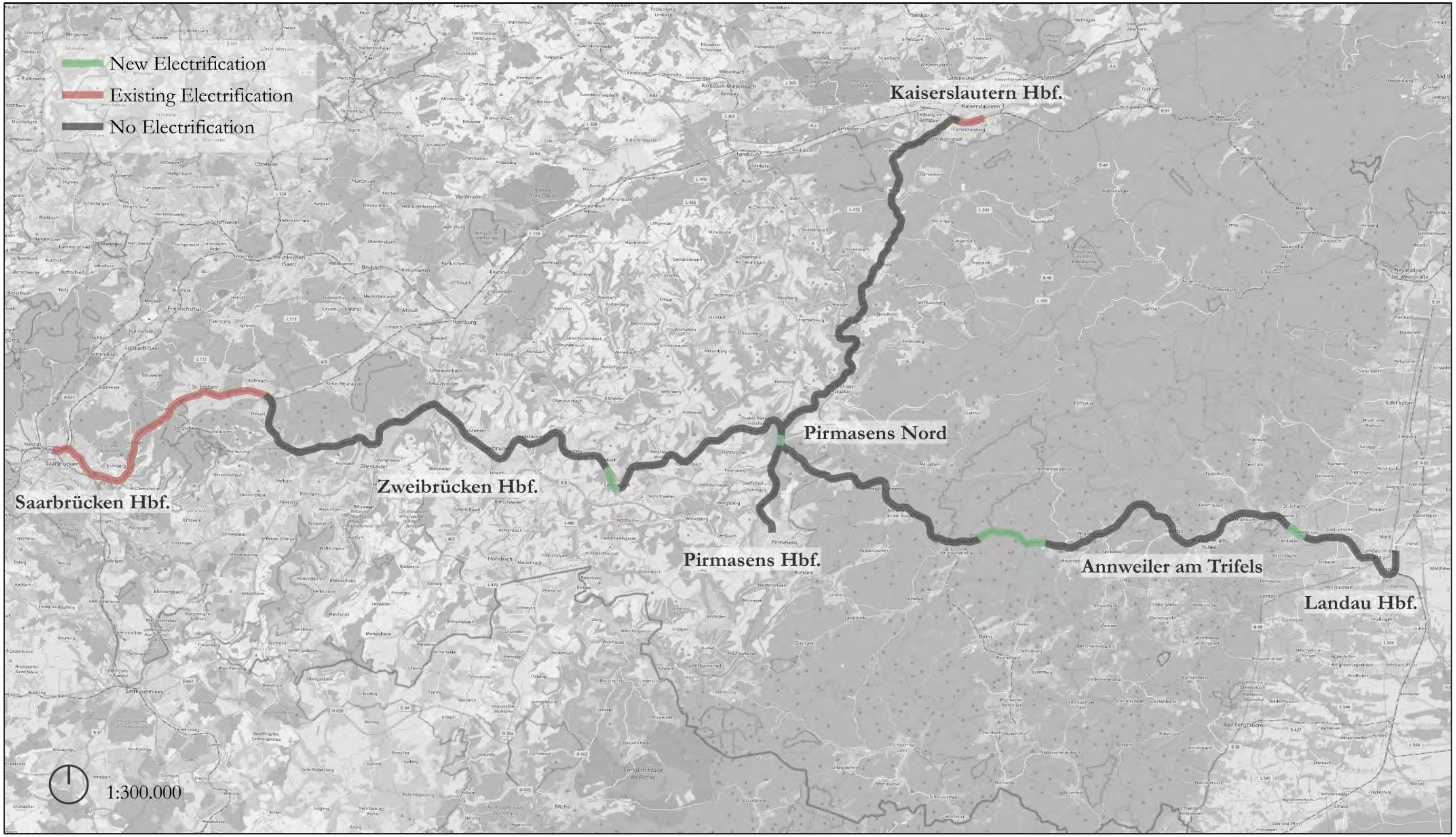

Figure 51: Existing and proposed electrification for model run "small battery 2". The electrification is given per track even though this is not visible at map scale.

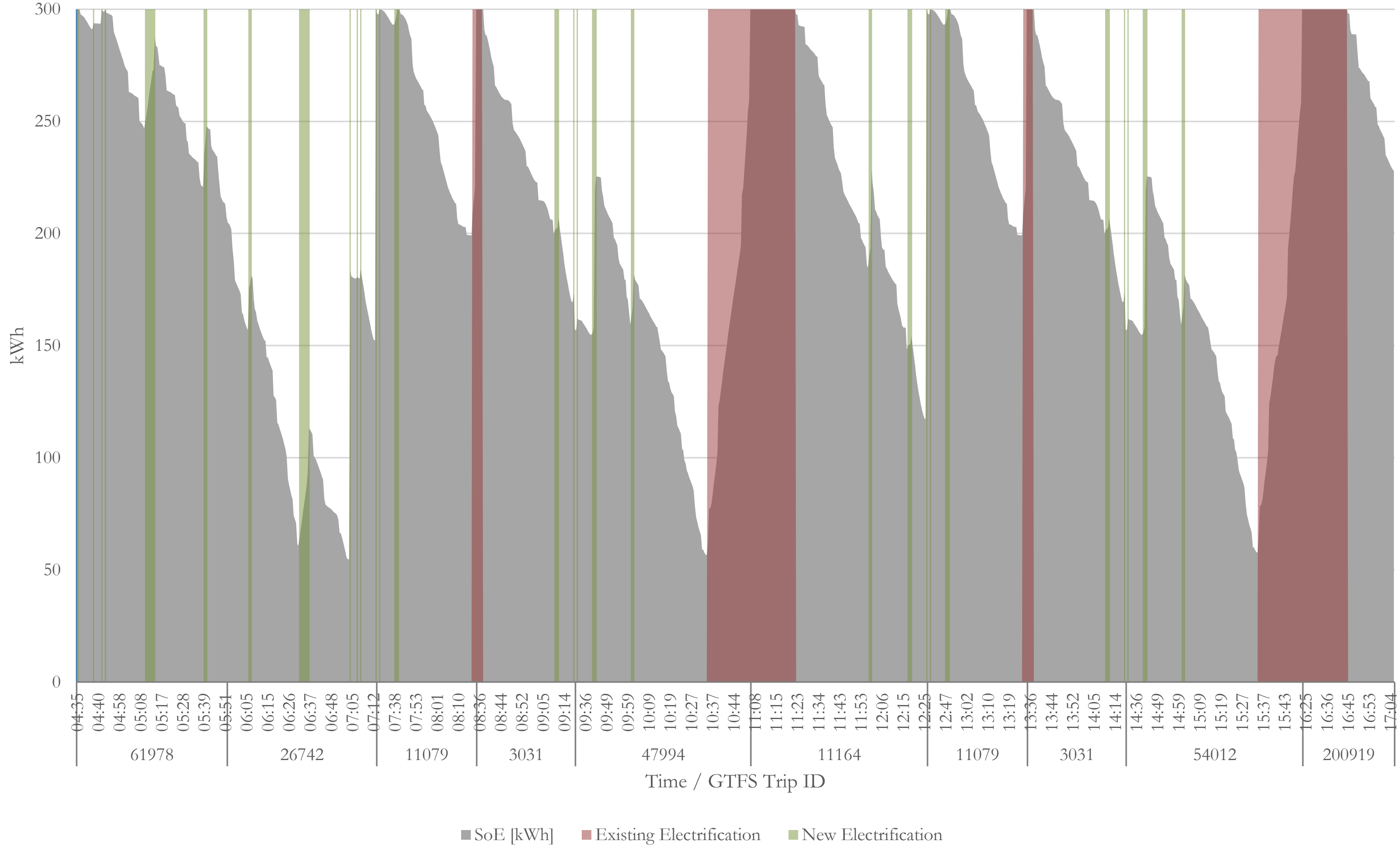

Figure 52: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run “small battery 2”.

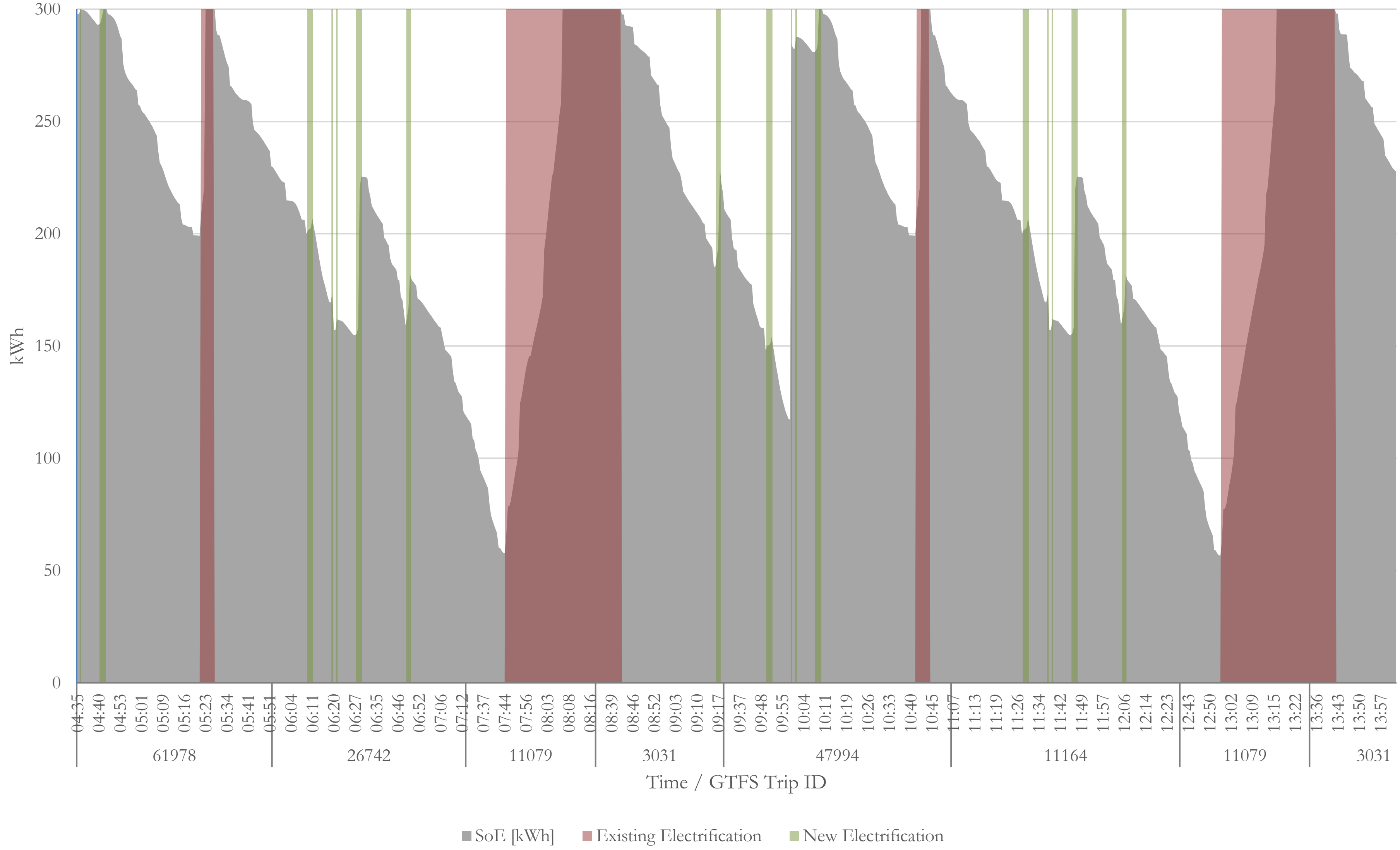

Figure 53: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run “small battery 2”.

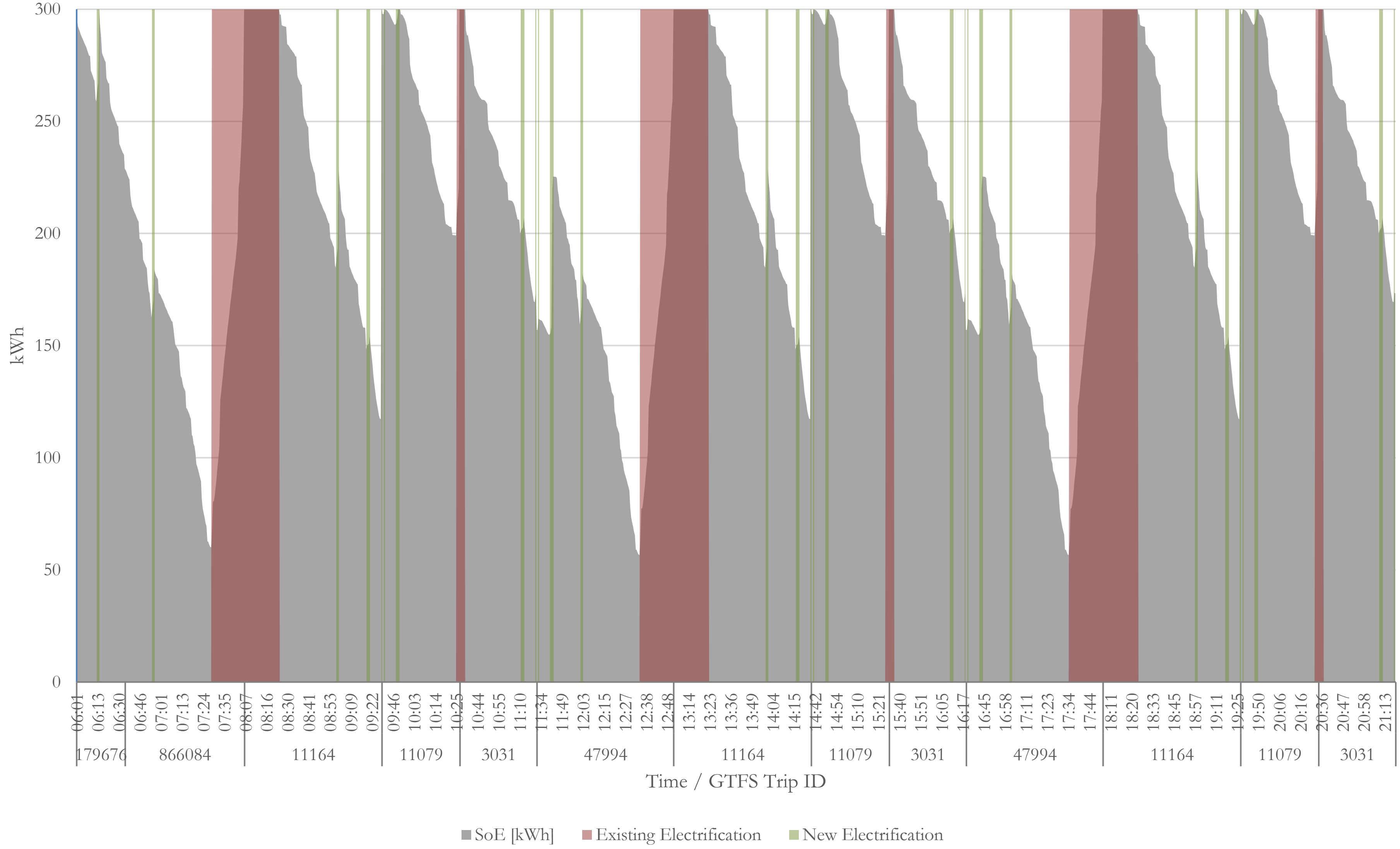

Figure 54: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run "small battery 2".

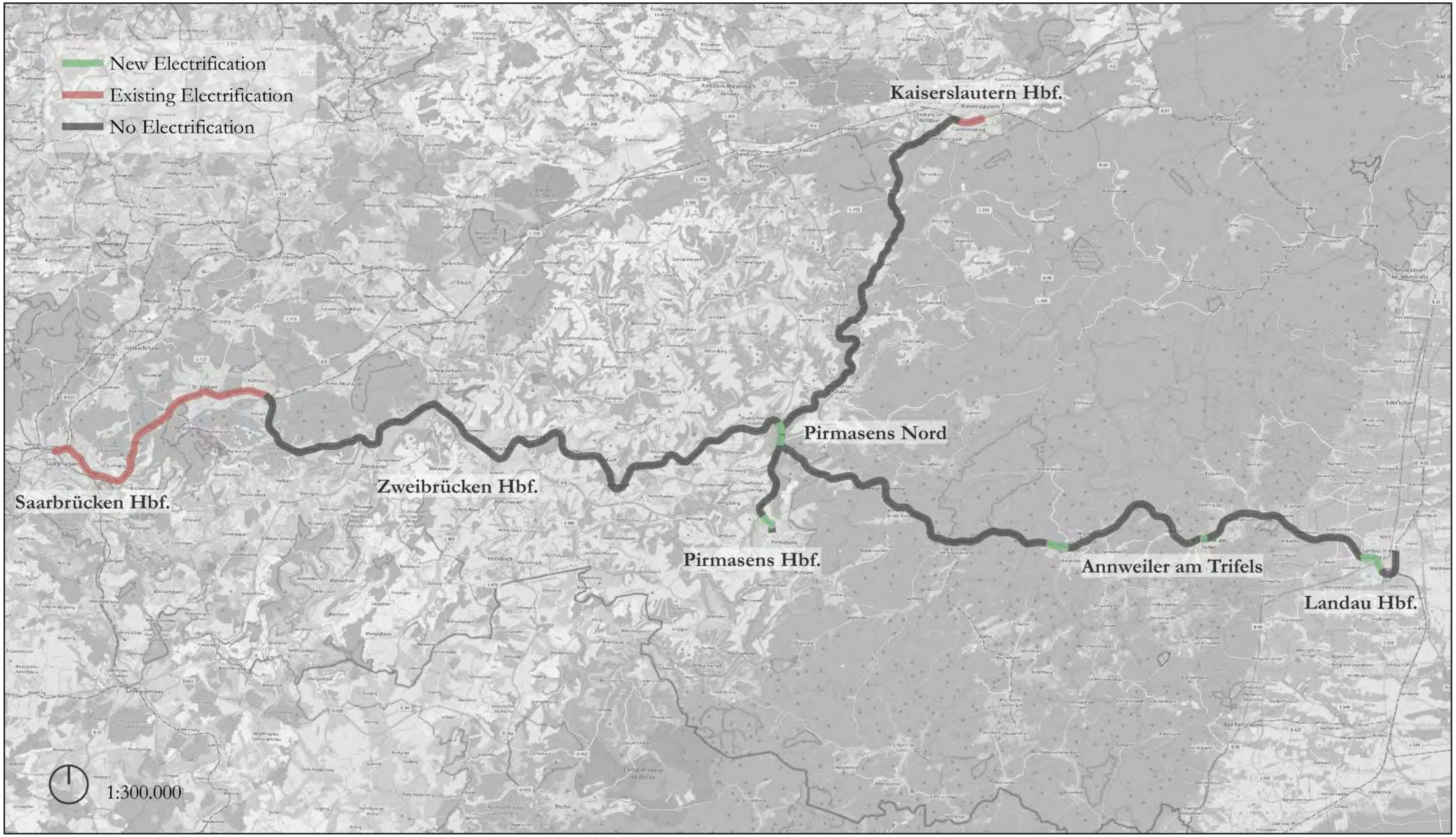

Figure 55: Existing and proposed electrification for model run “small battery 3”. The electrification is given per track even though this is not visible at map scale.

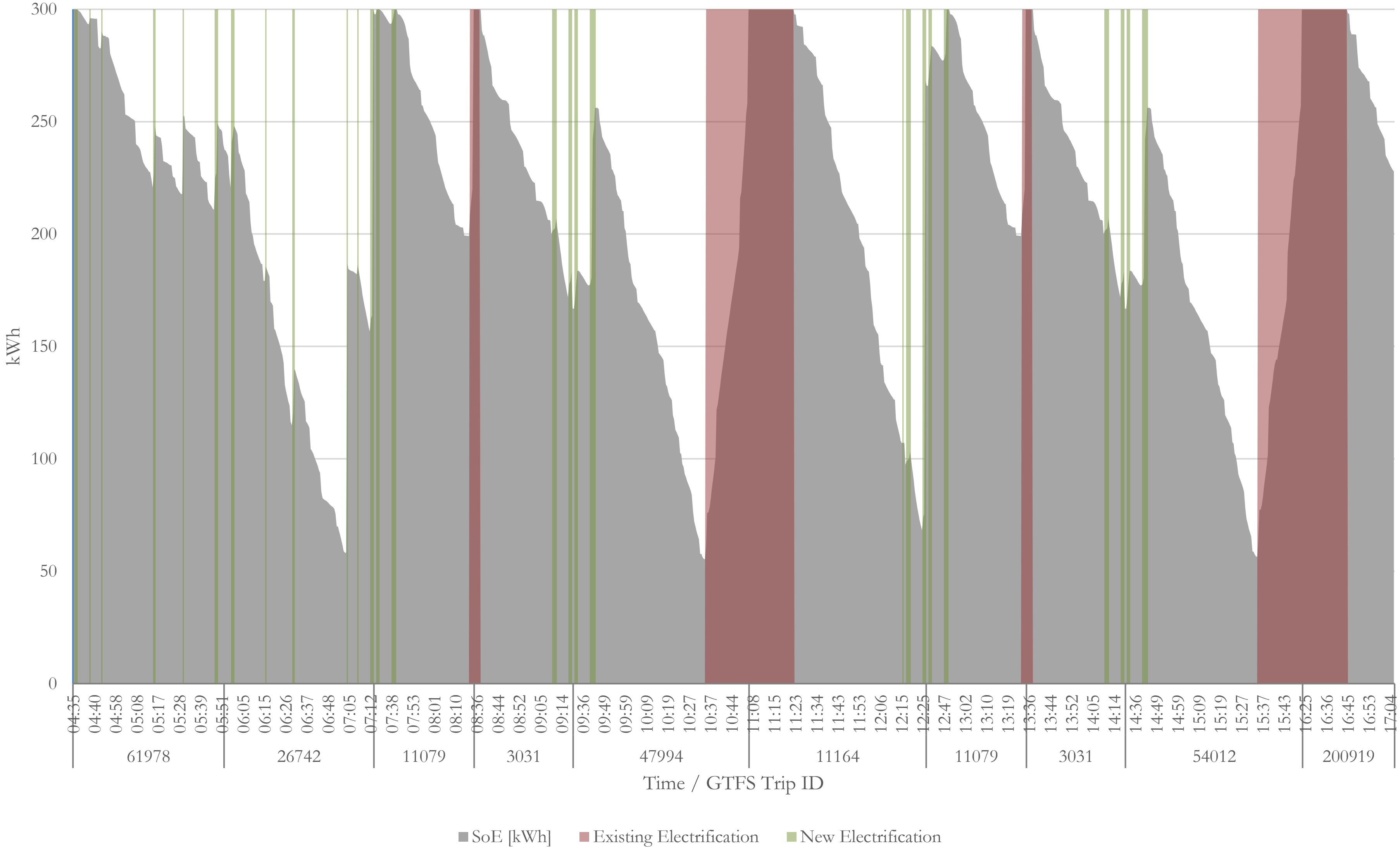

Figure 56: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run "small battery 3".

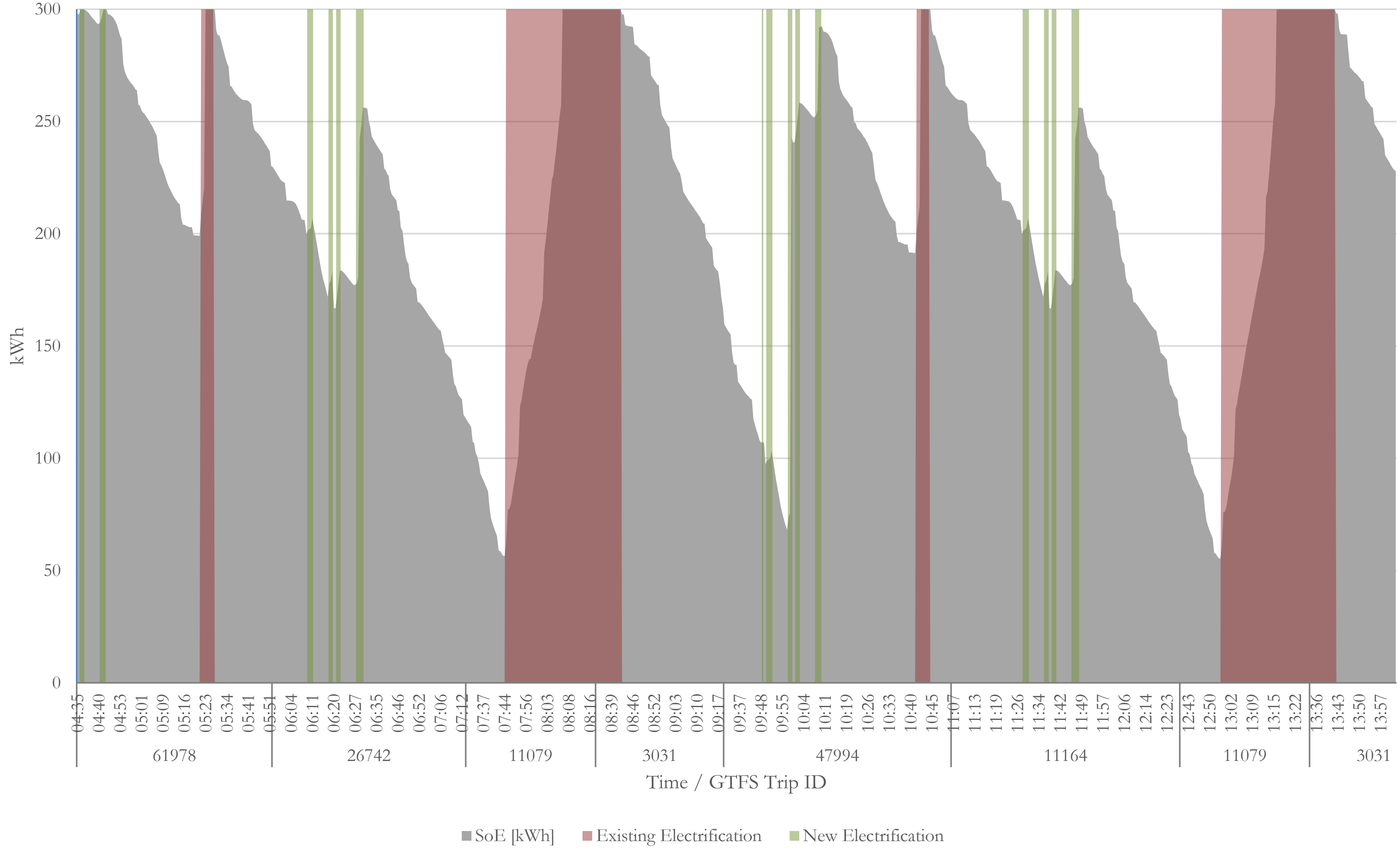

Figure 57: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run "small battery 3".

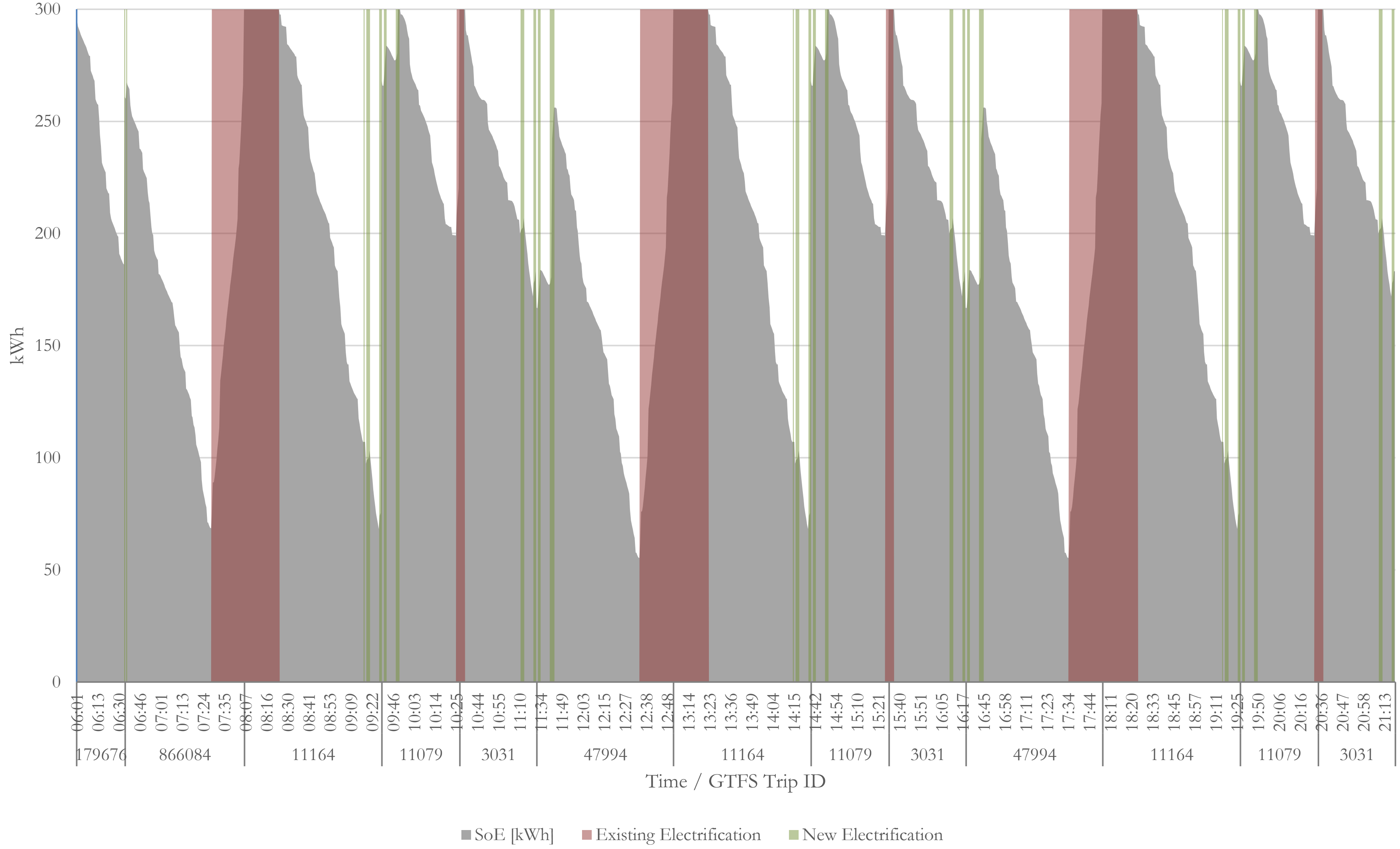

Figure 58: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run “small battery 3”.

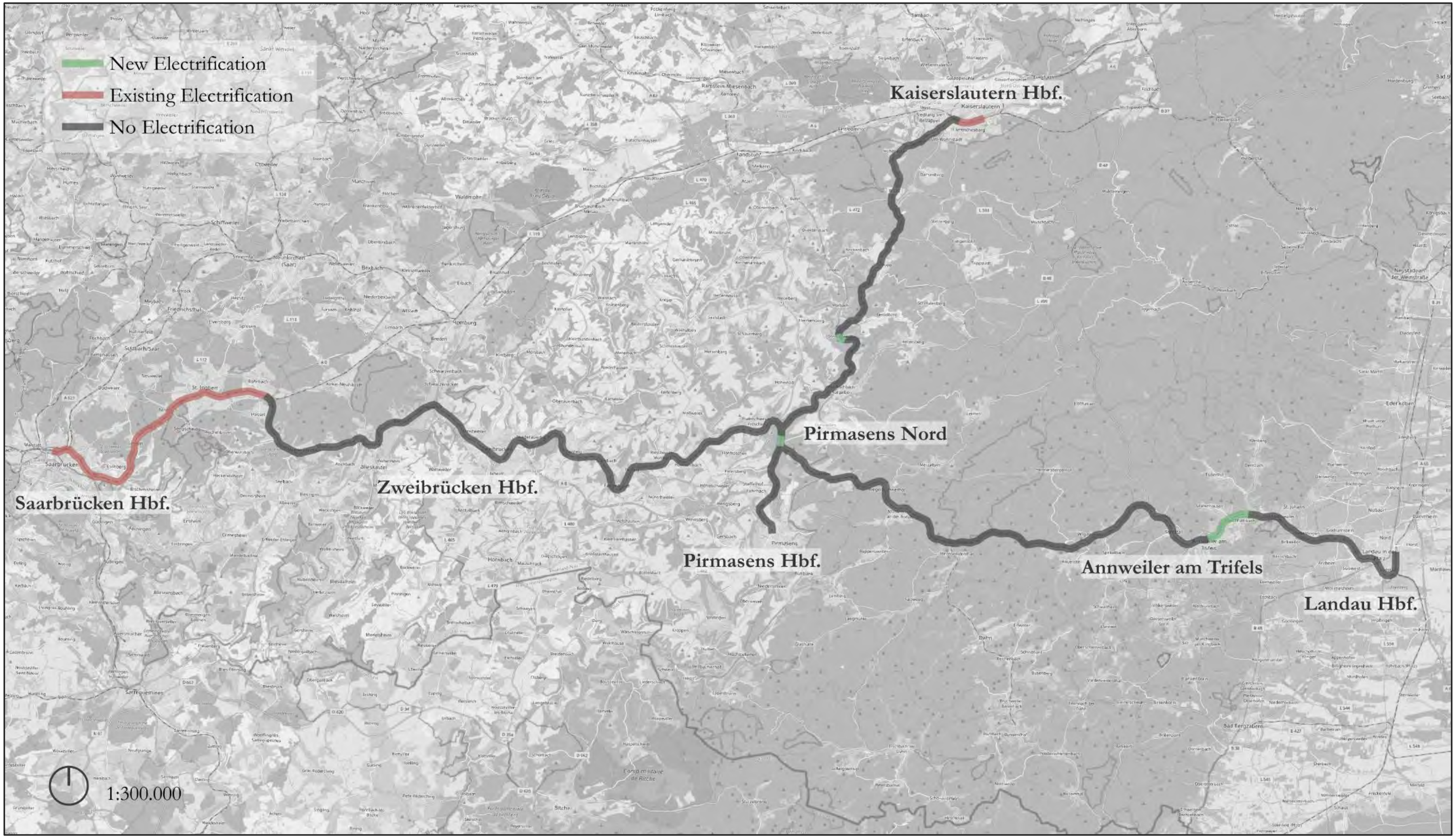

Figure 59: Existing and proposed electrification for model run “standard battery 1”. The electrification is given per track even though this is not visible at map scale.

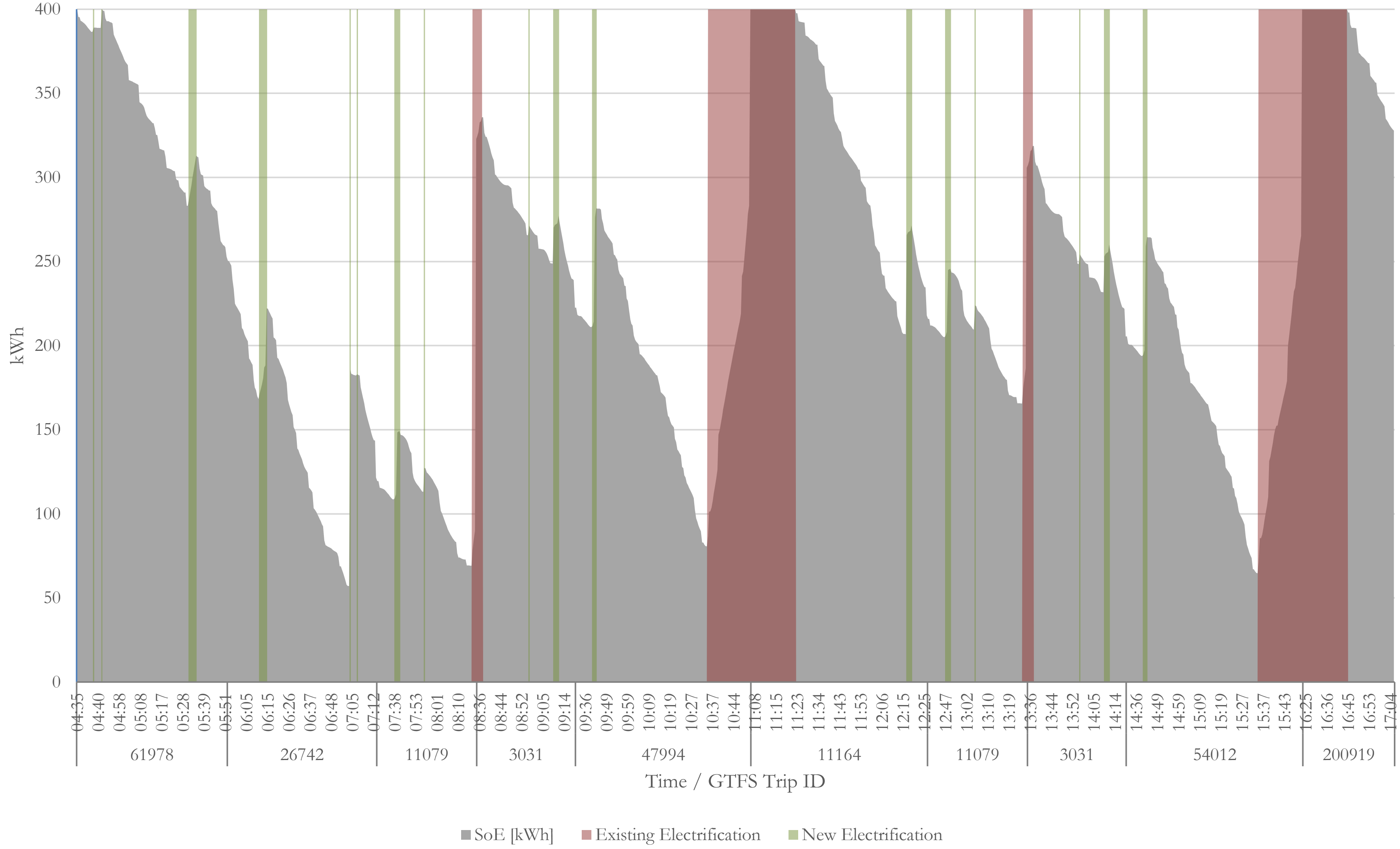

Figure 60: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run “standard battery 1”.

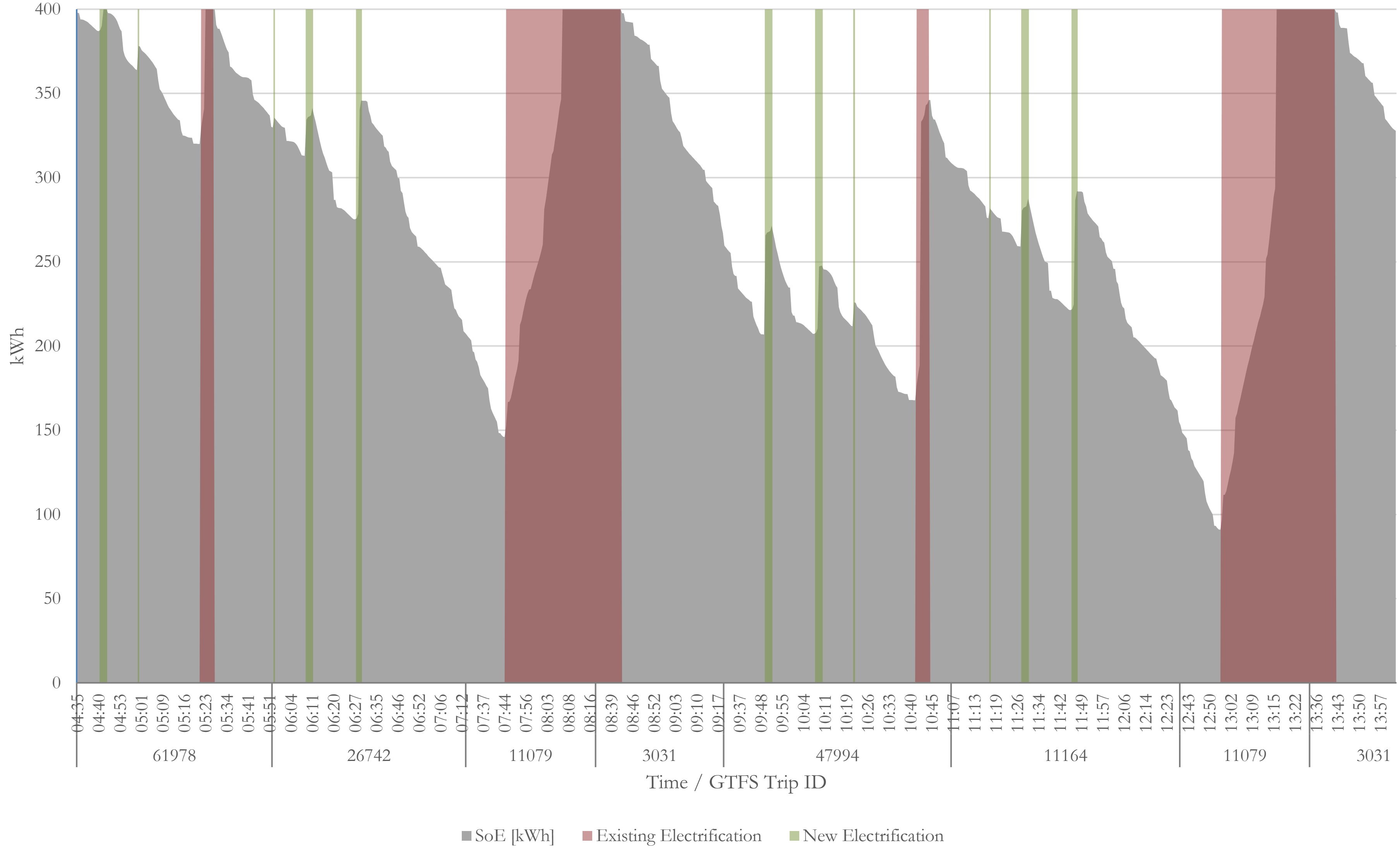

Figure 61: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run "standard battery 1".

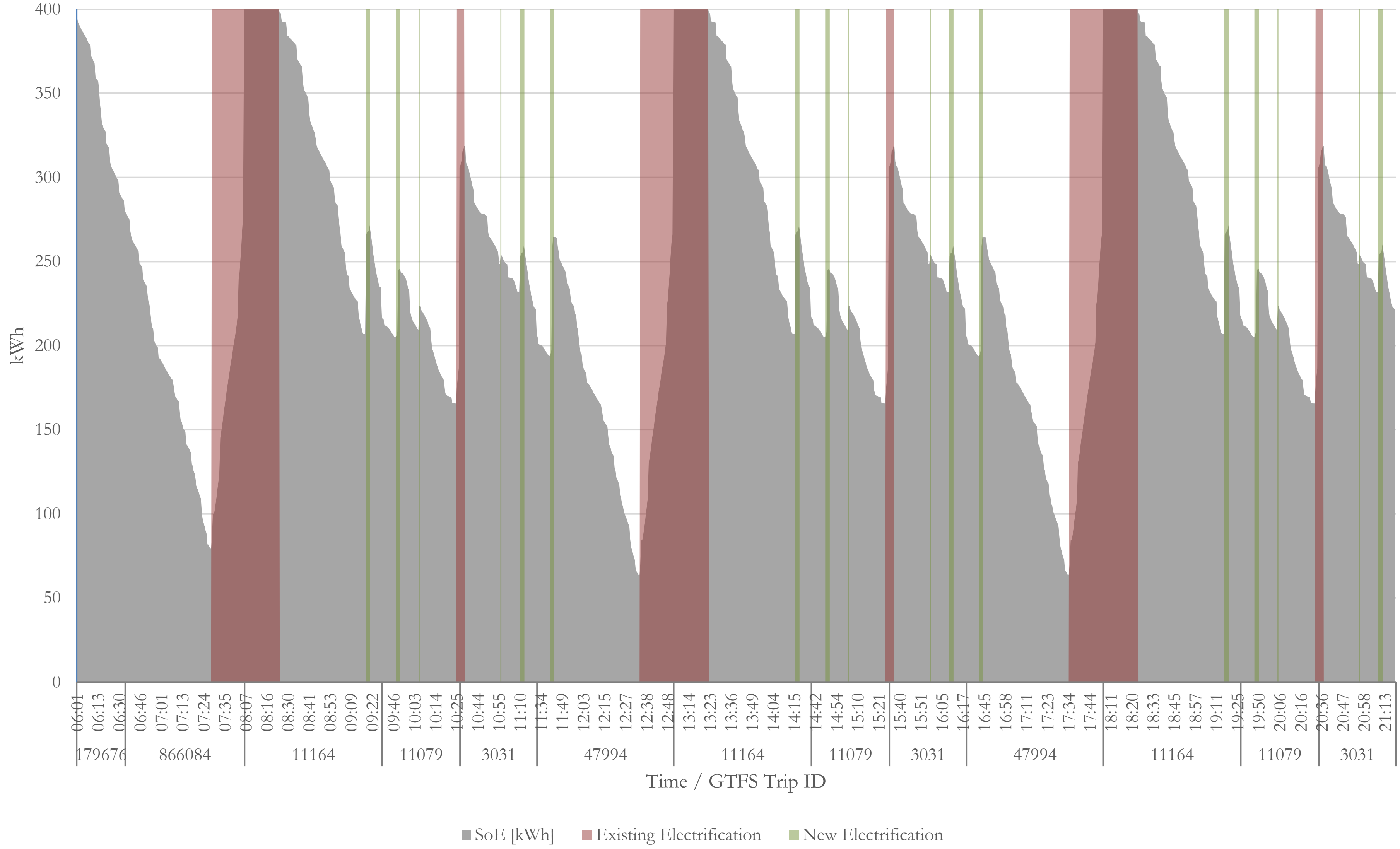

Figure 62: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run “standard battery 1”.

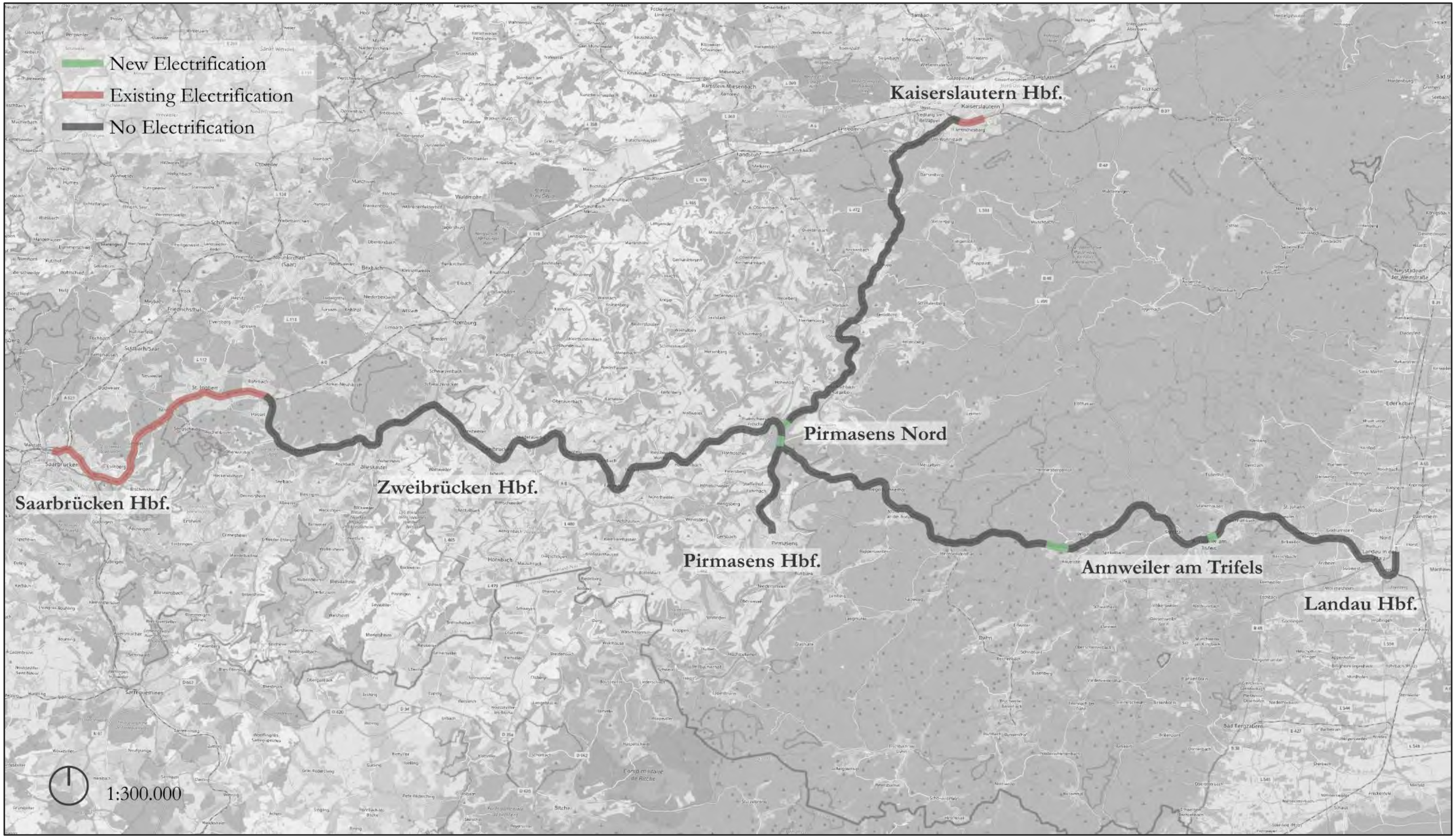

Figure 63: Existing and proposed electrification for model run "standard battery 2". The electrification is given per track even though this is not visible at map scale.

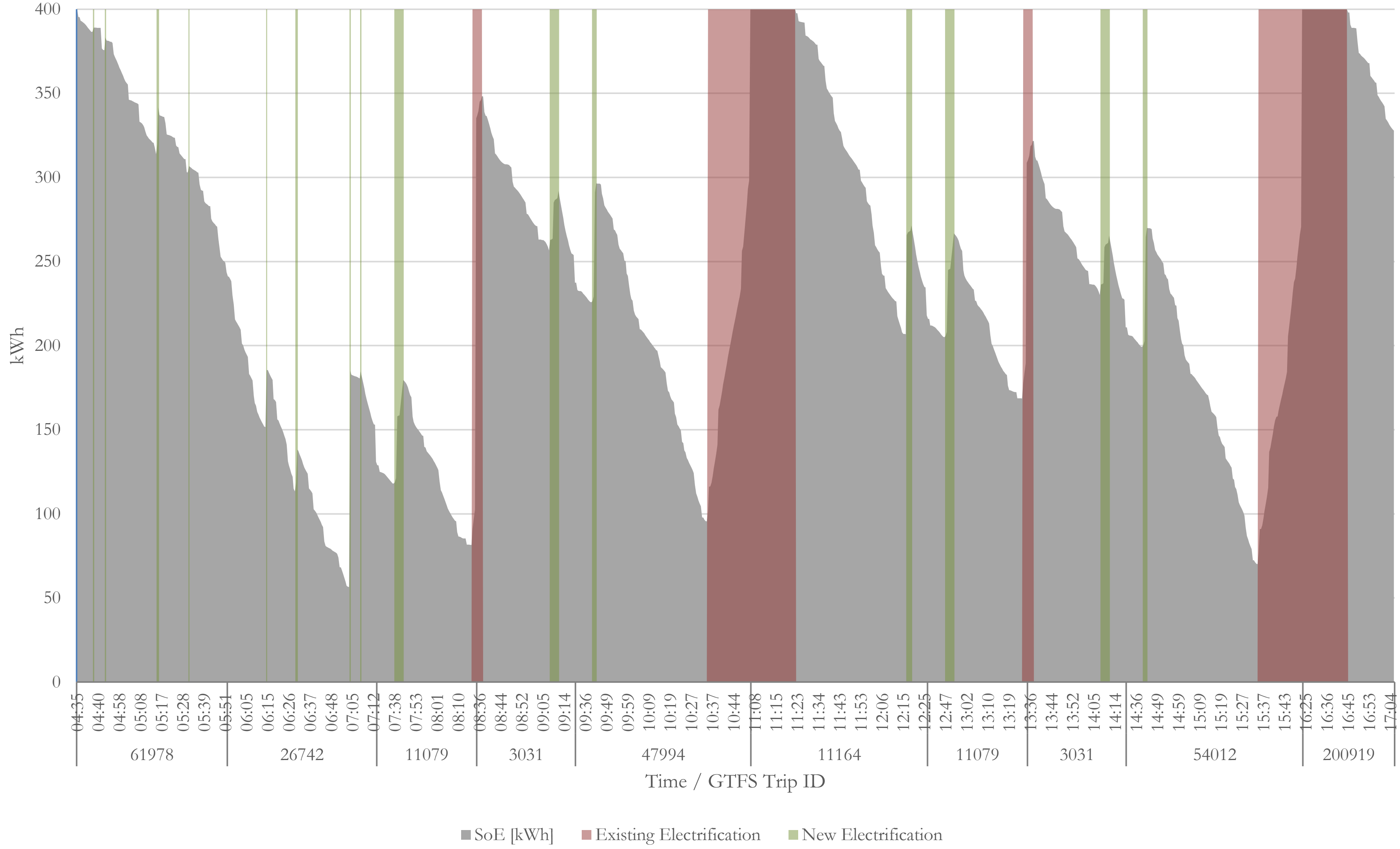

Figure 64: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run "standard battery 2".

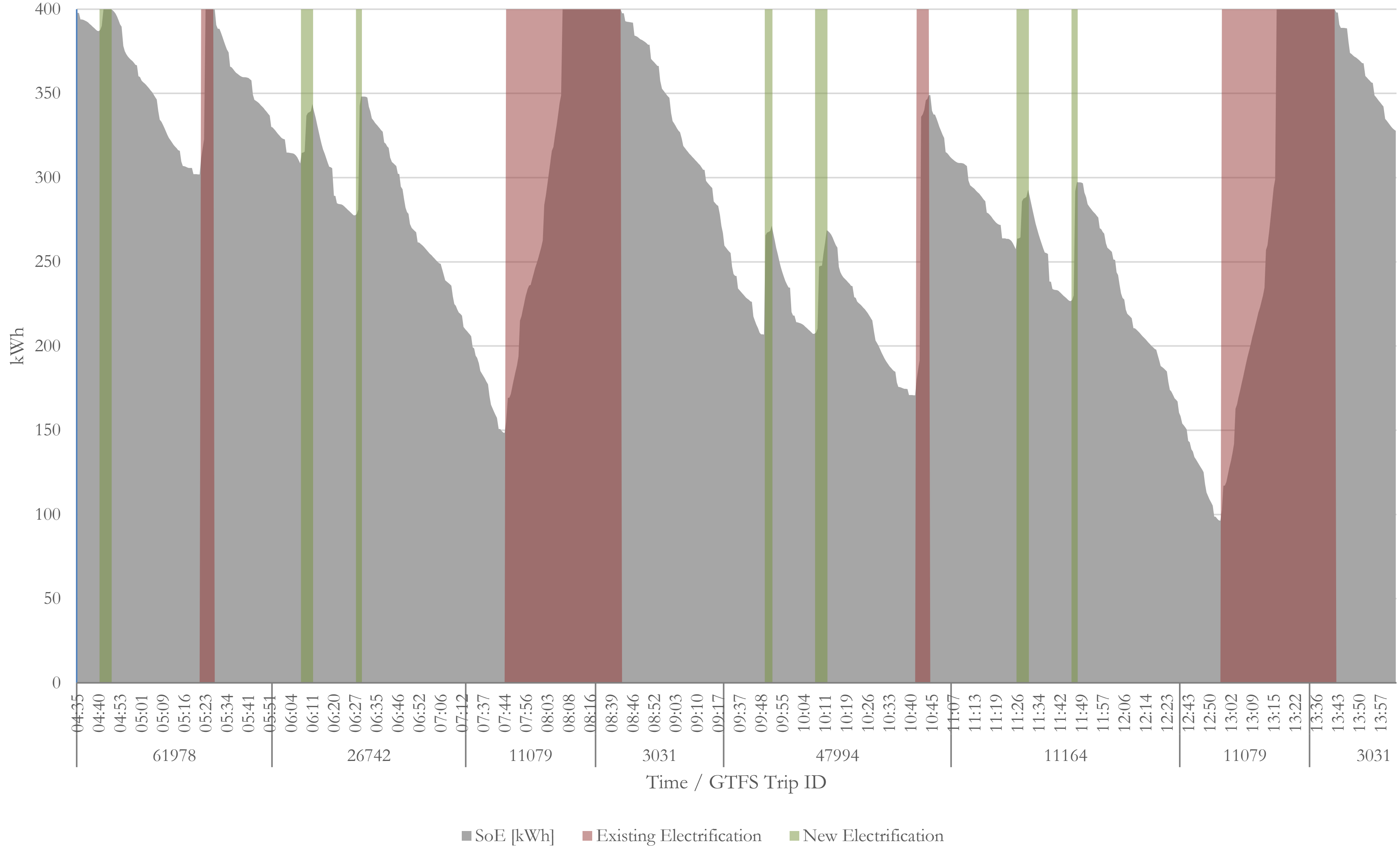

Figure 65: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run "standard battery 2".

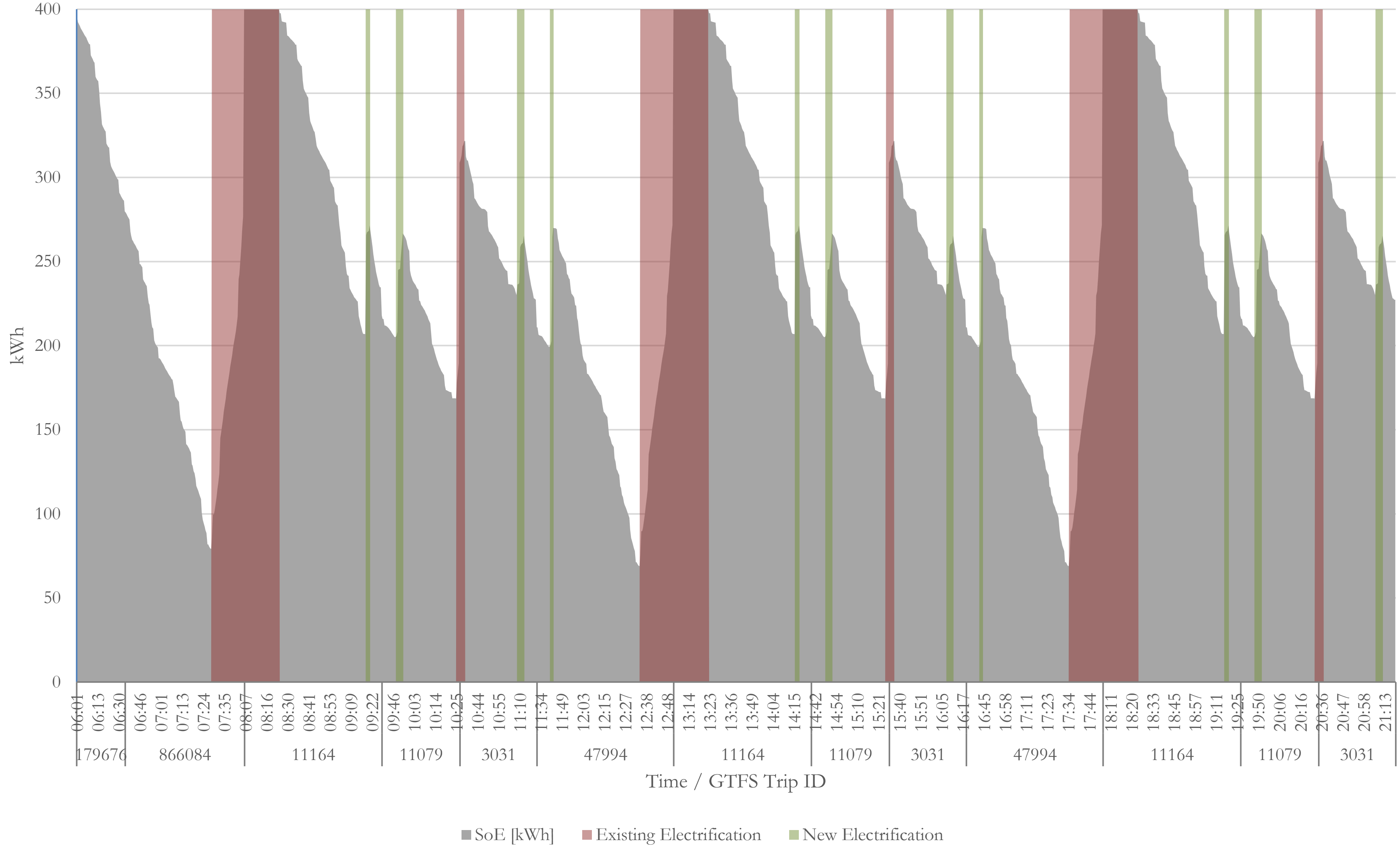

Figure 66: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run “standard battery 2”.

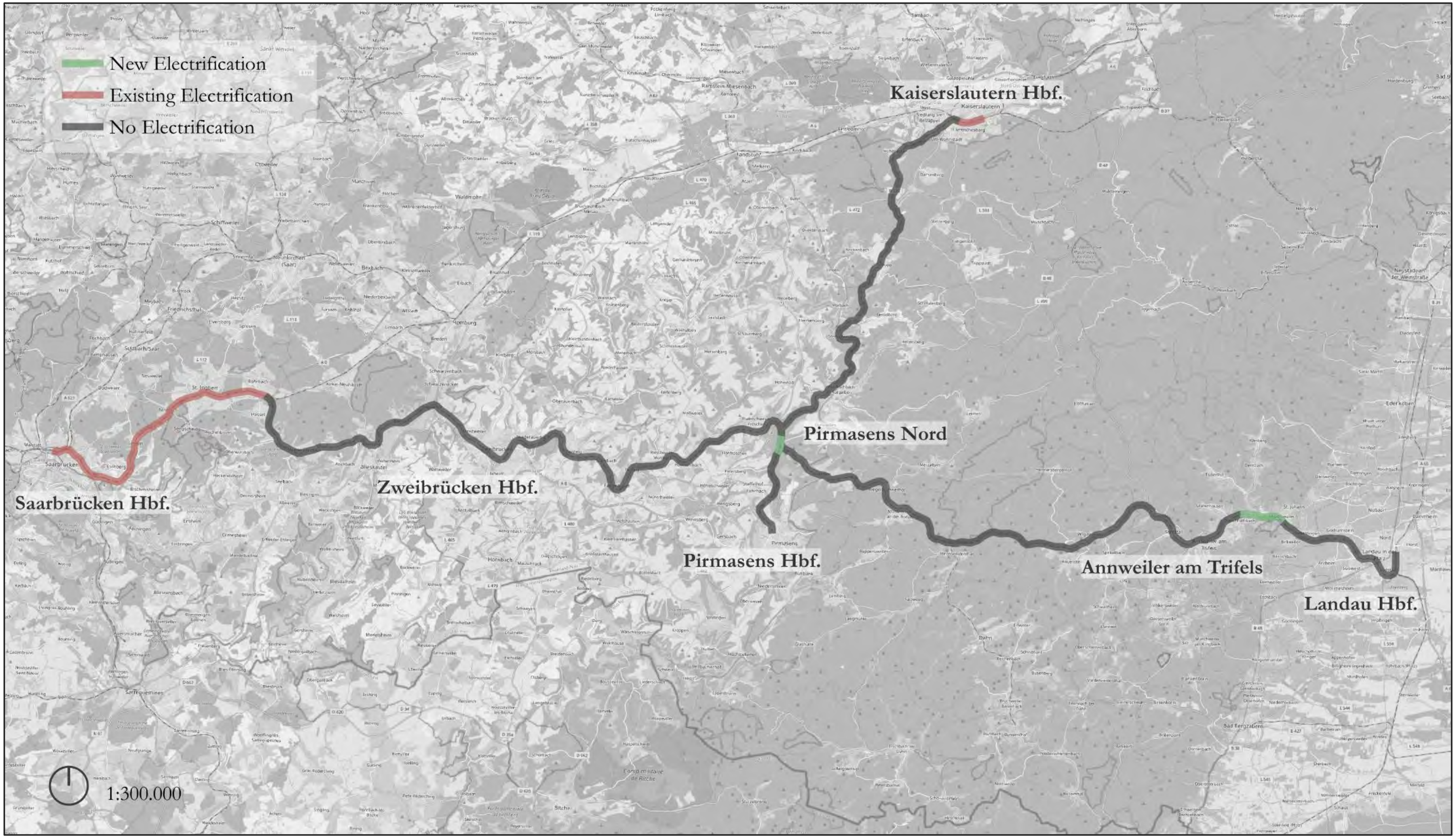

Figure 67: Existing and proposed electrification for model run “standard battery 3”. The electrification is given per track even though this is not visible at map scale.

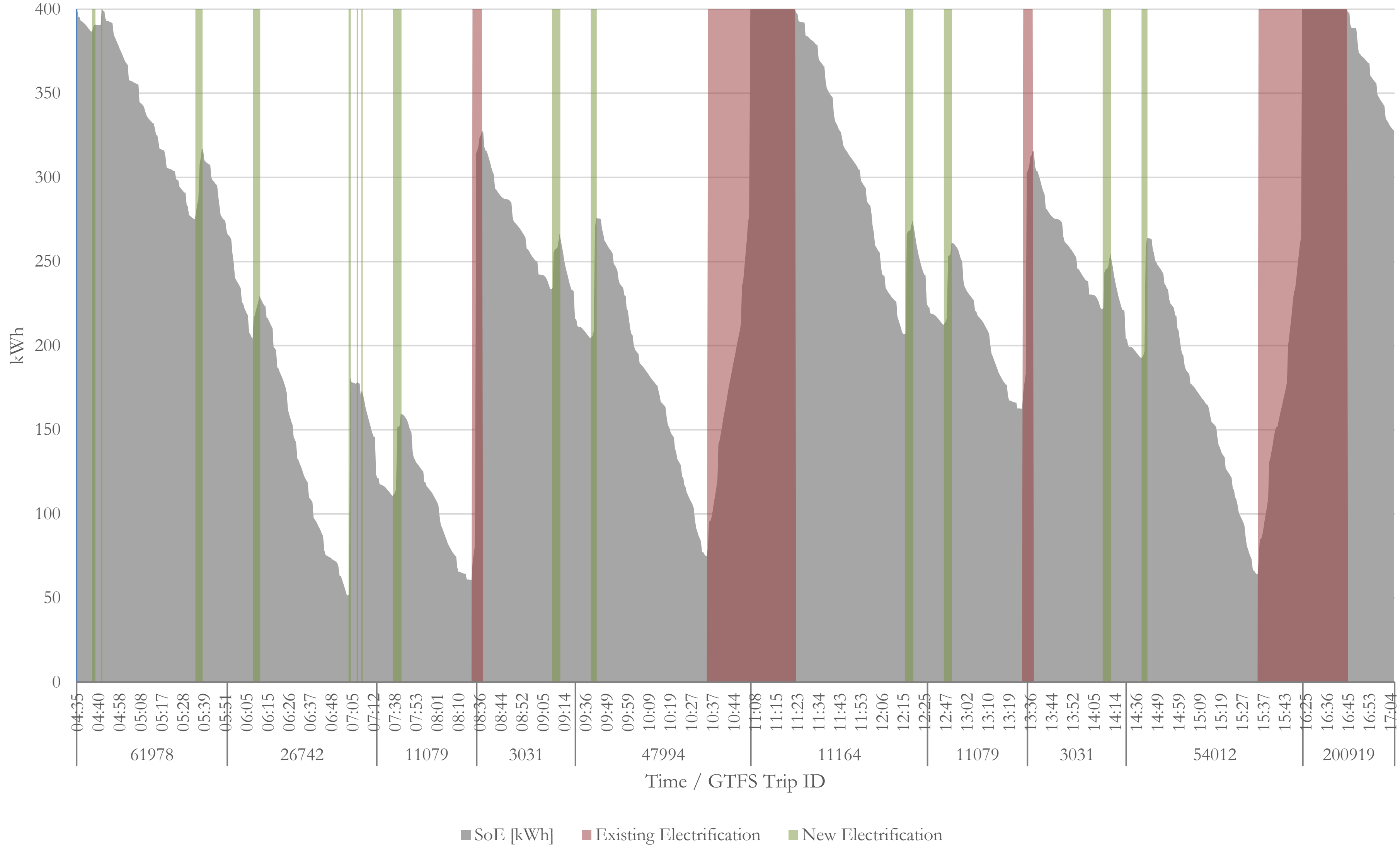

Figure 68: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run “standard battery 3”.

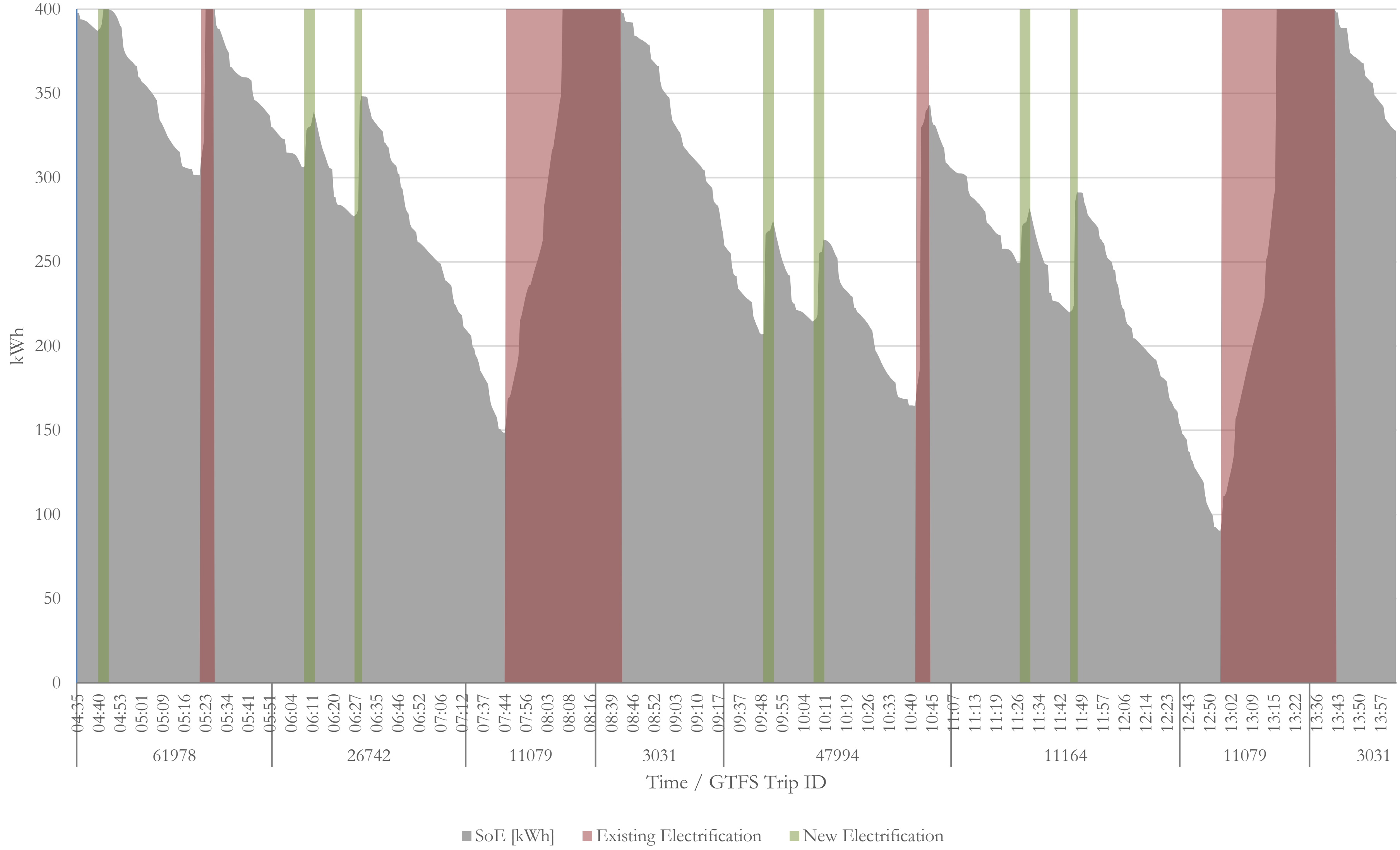

Figure 69: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run “standard battery 3”.

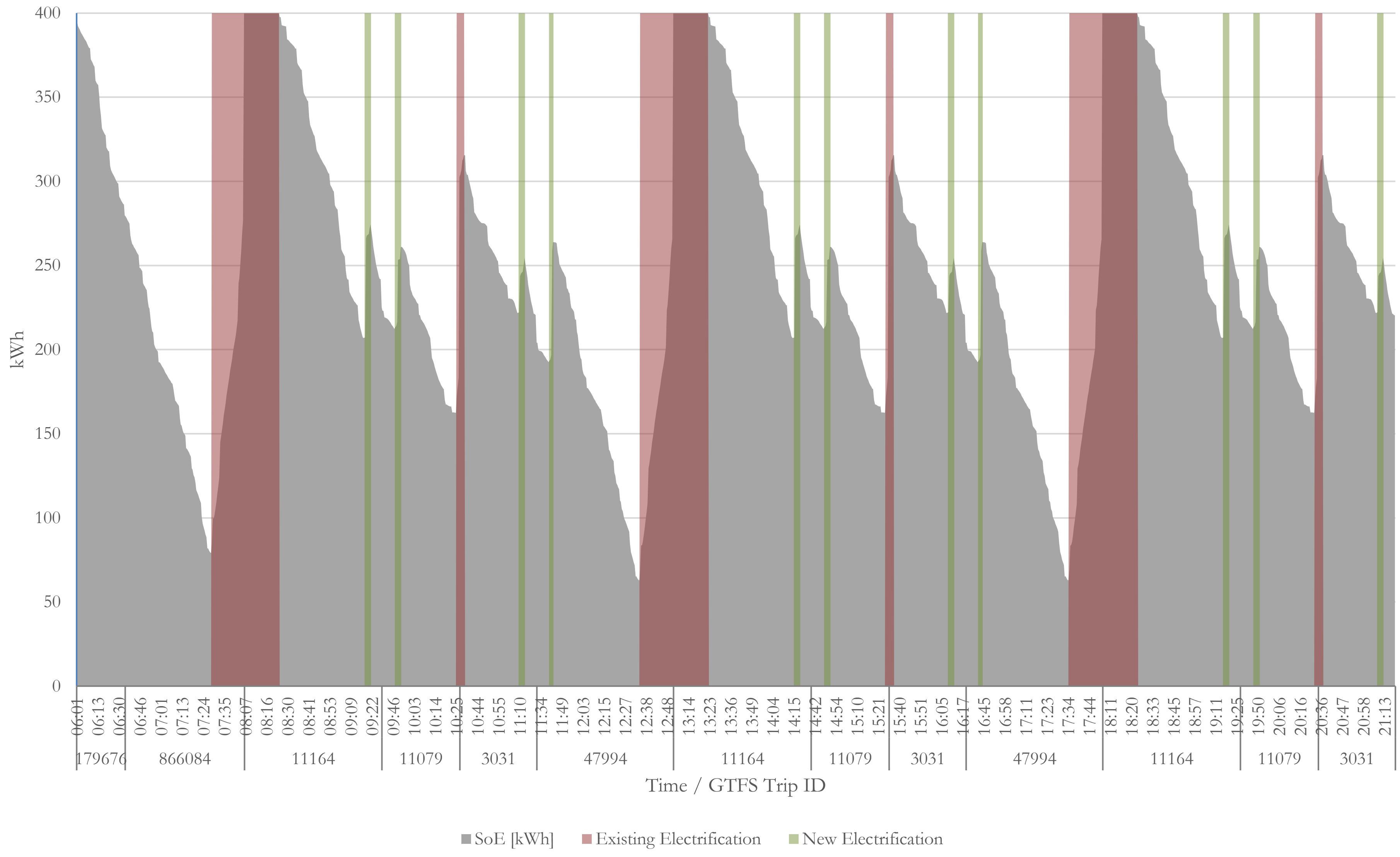

Figure 70: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run "standard battery 3".

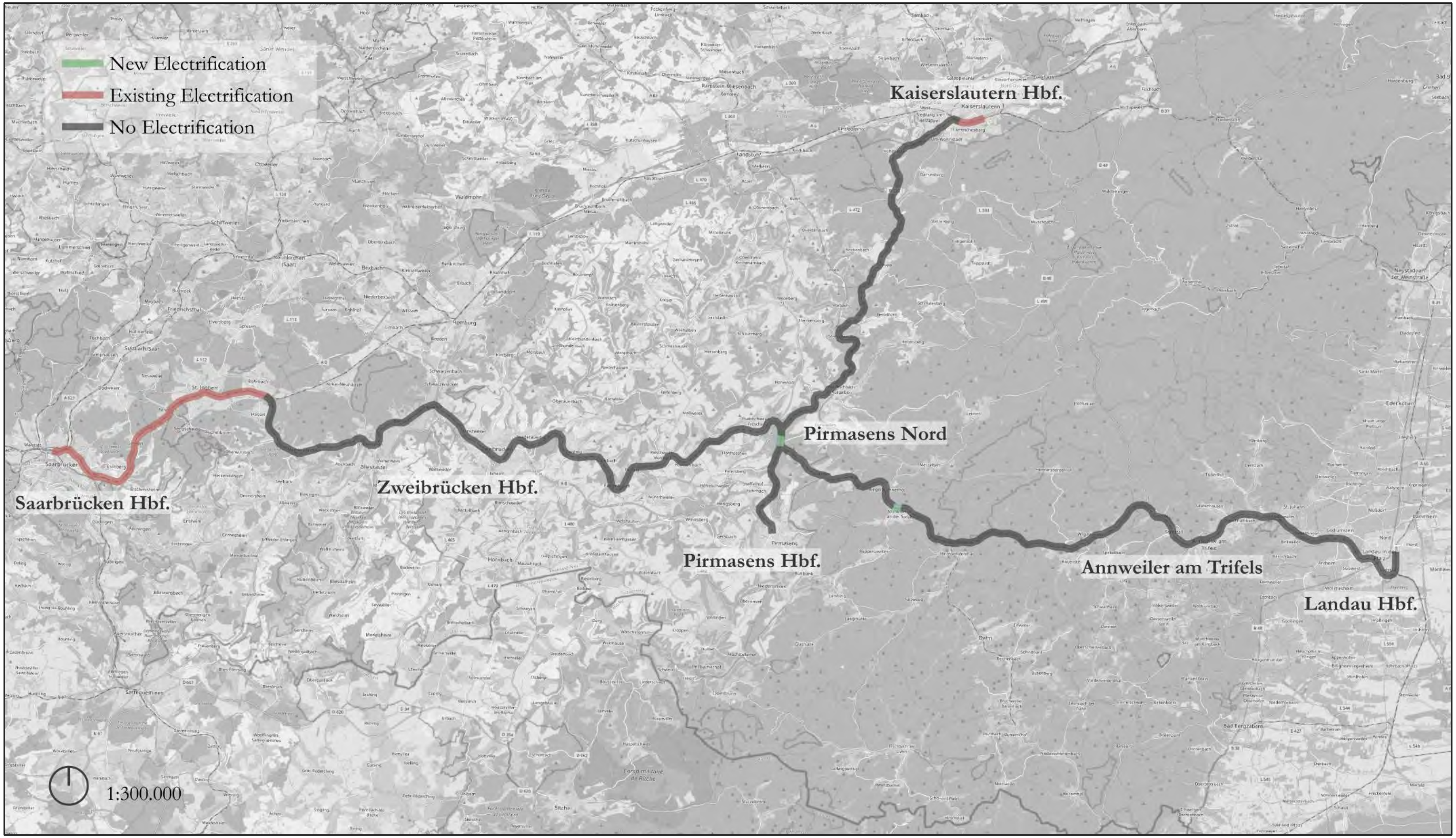

Figure 71: Existing and proposed electrification for model run "large battery 1". The electrification is given per track even though this is not visible at map scale.

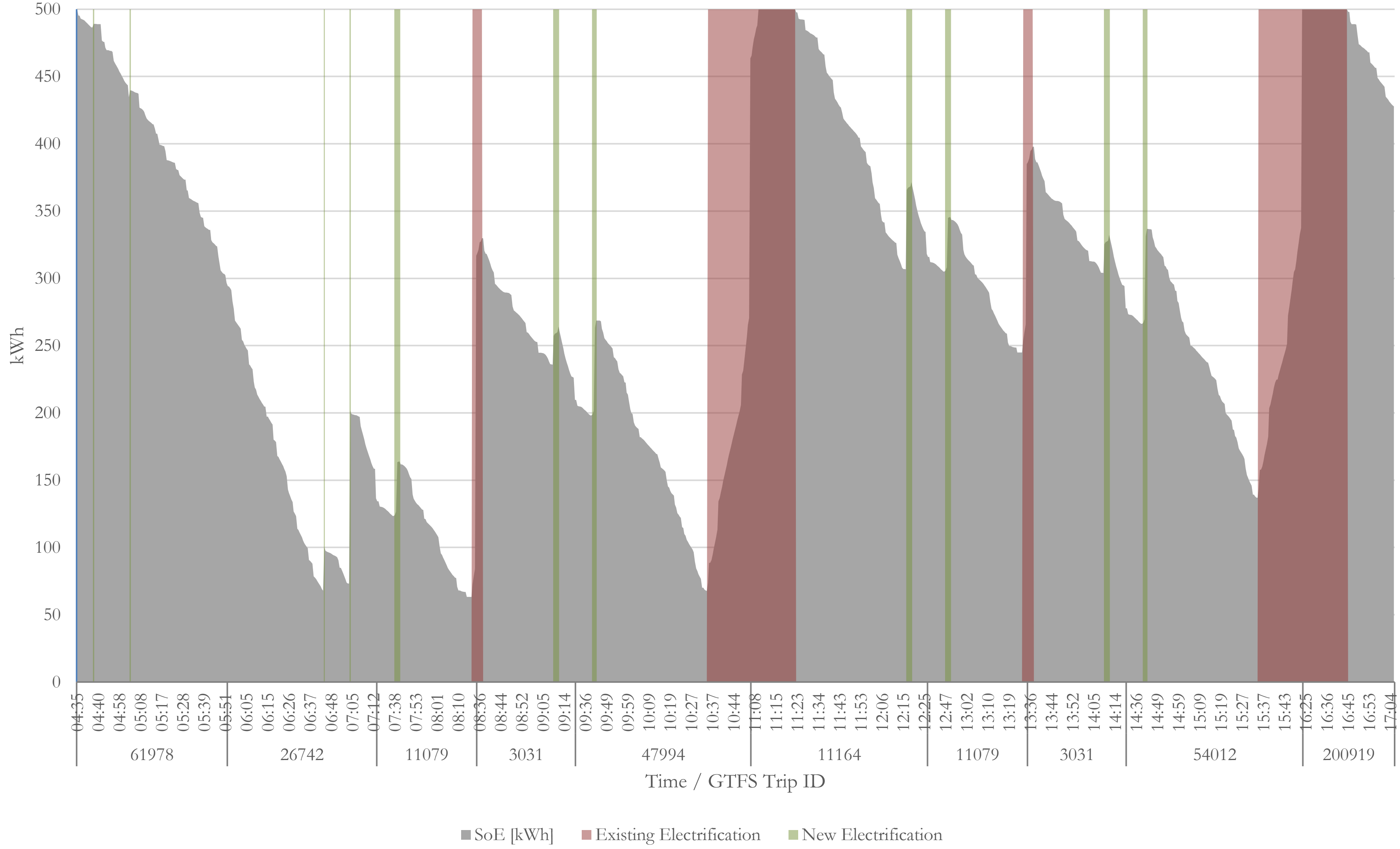

Figure 72: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run “large battery 1”.

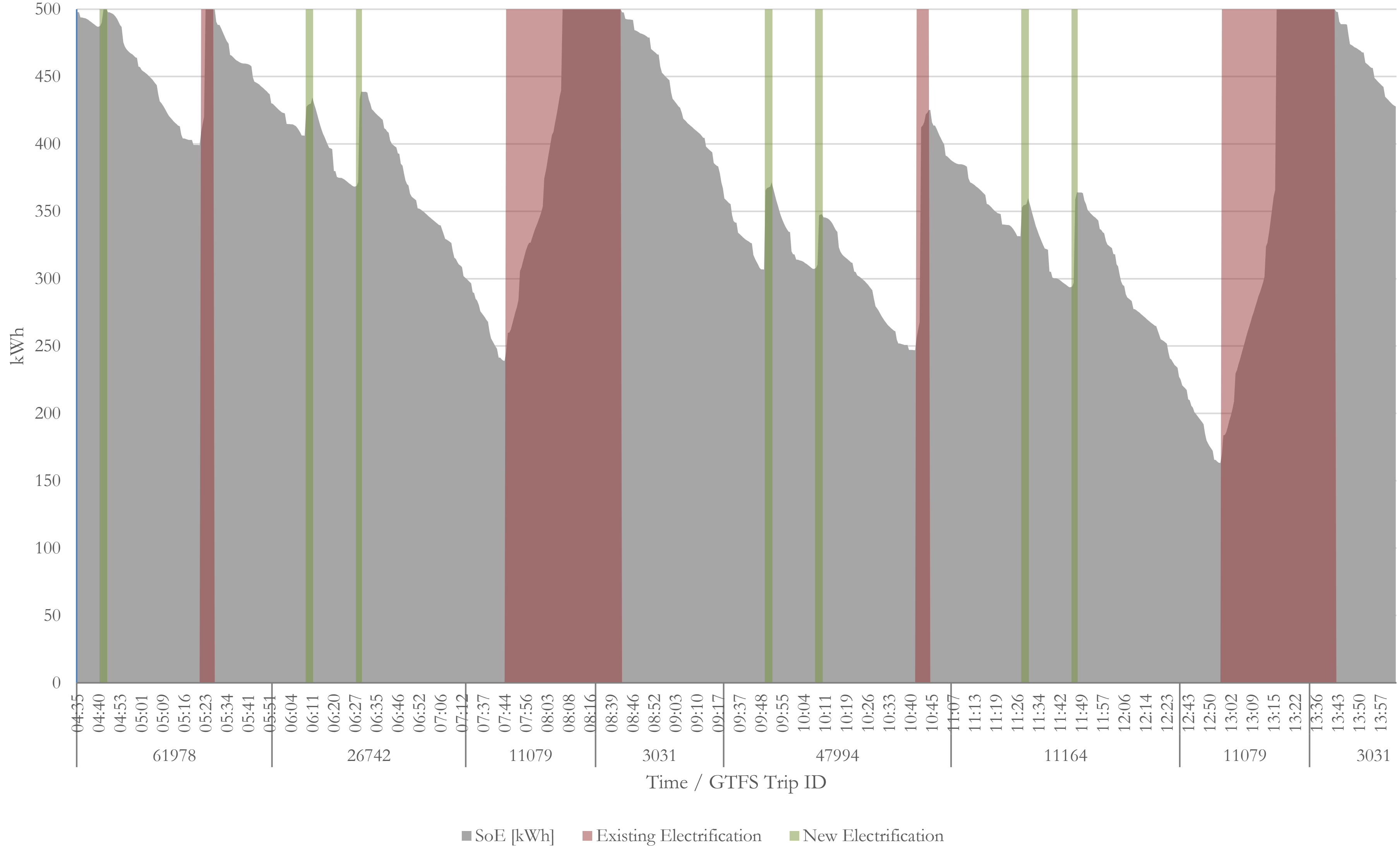

Figure 73: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run "large battery 1".

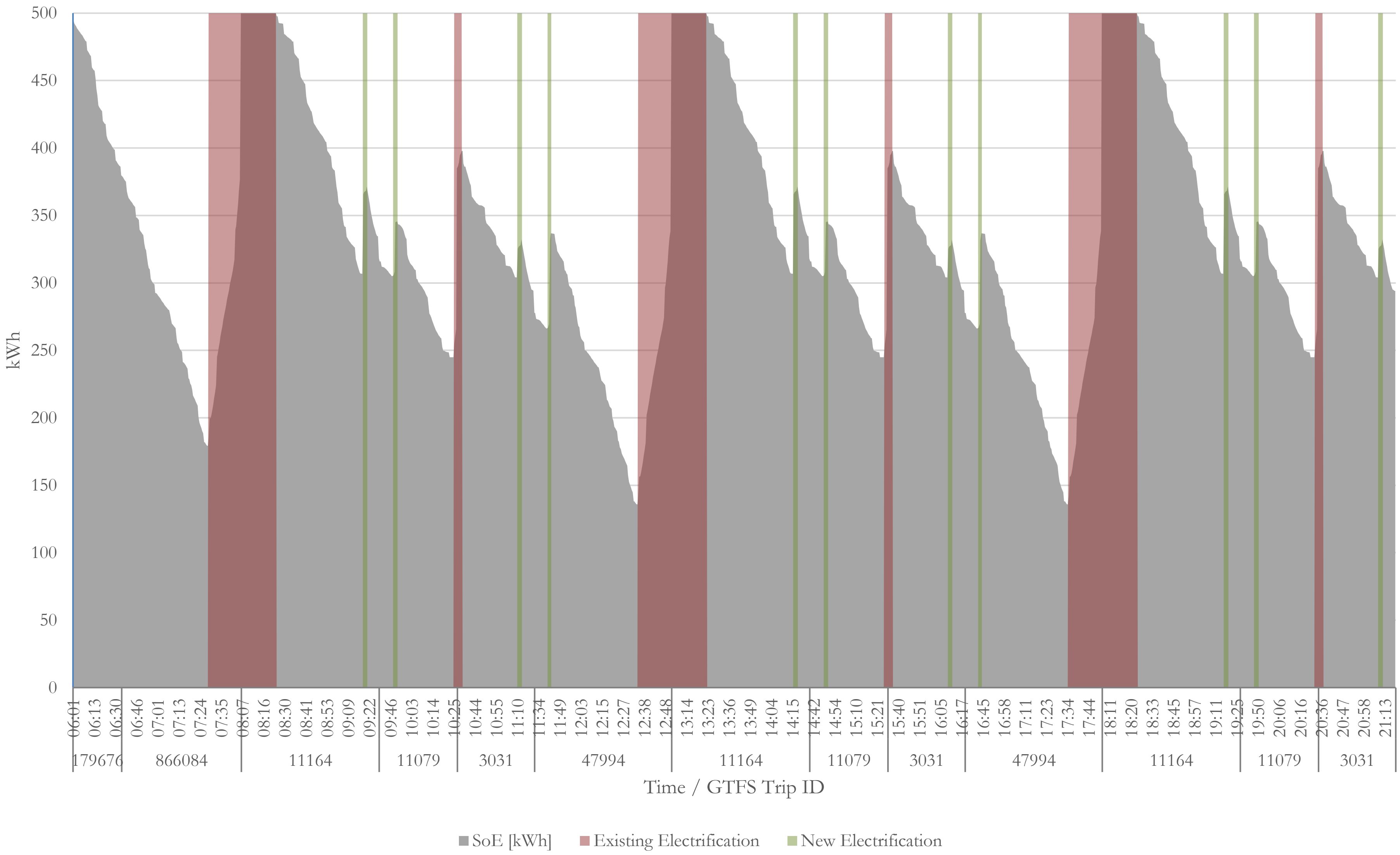

Figure 74: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run "large battery 1".

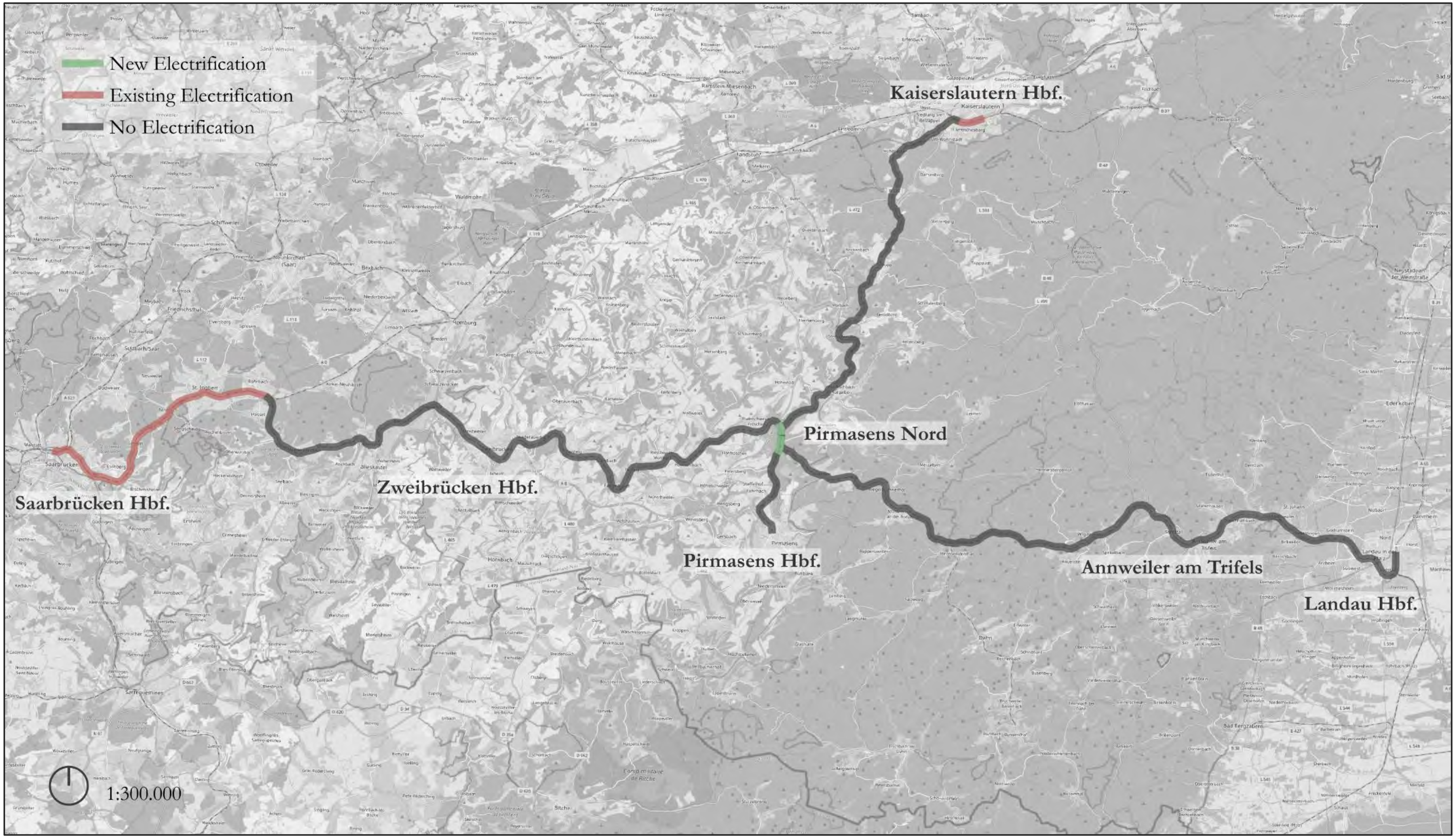

Figure 75: Existing and proposed electrification for model run “large battery 2”. The electrification is given per track even though this is not visible at map scale.

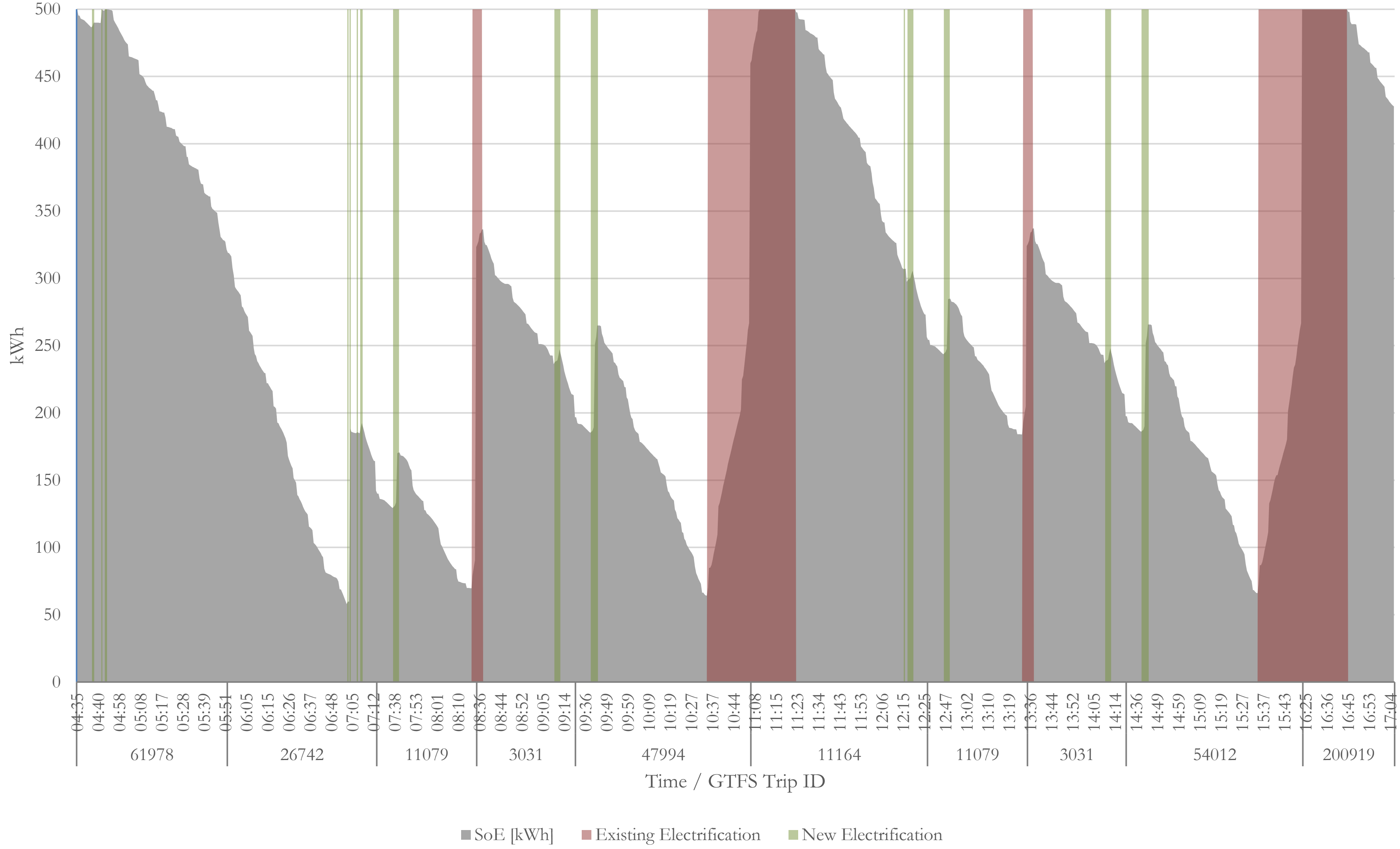

Figure 76: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run "large battery 2".

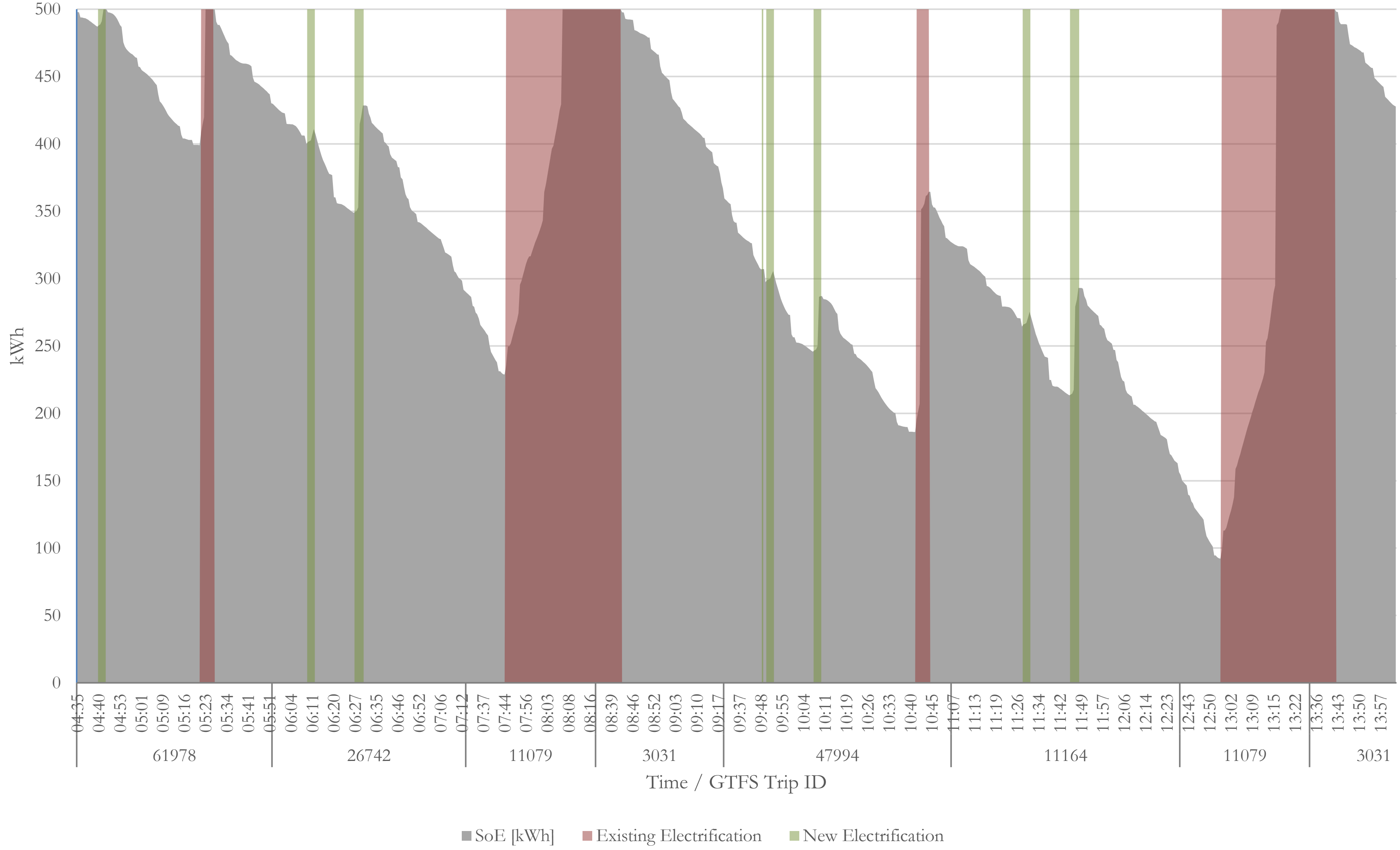

Figure 77: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run “large battery 2”.

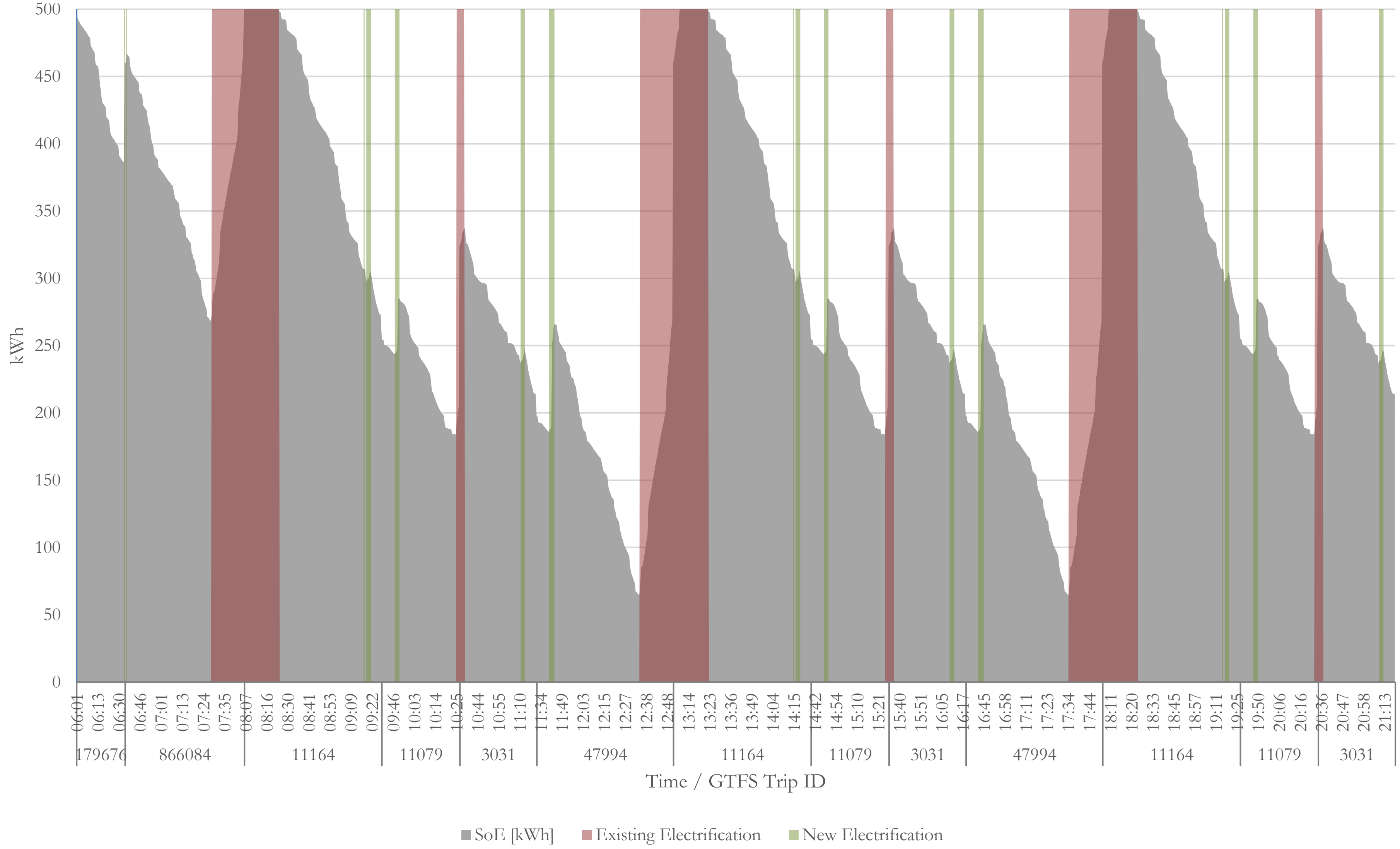

Figure 78: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run “large battery 2”.

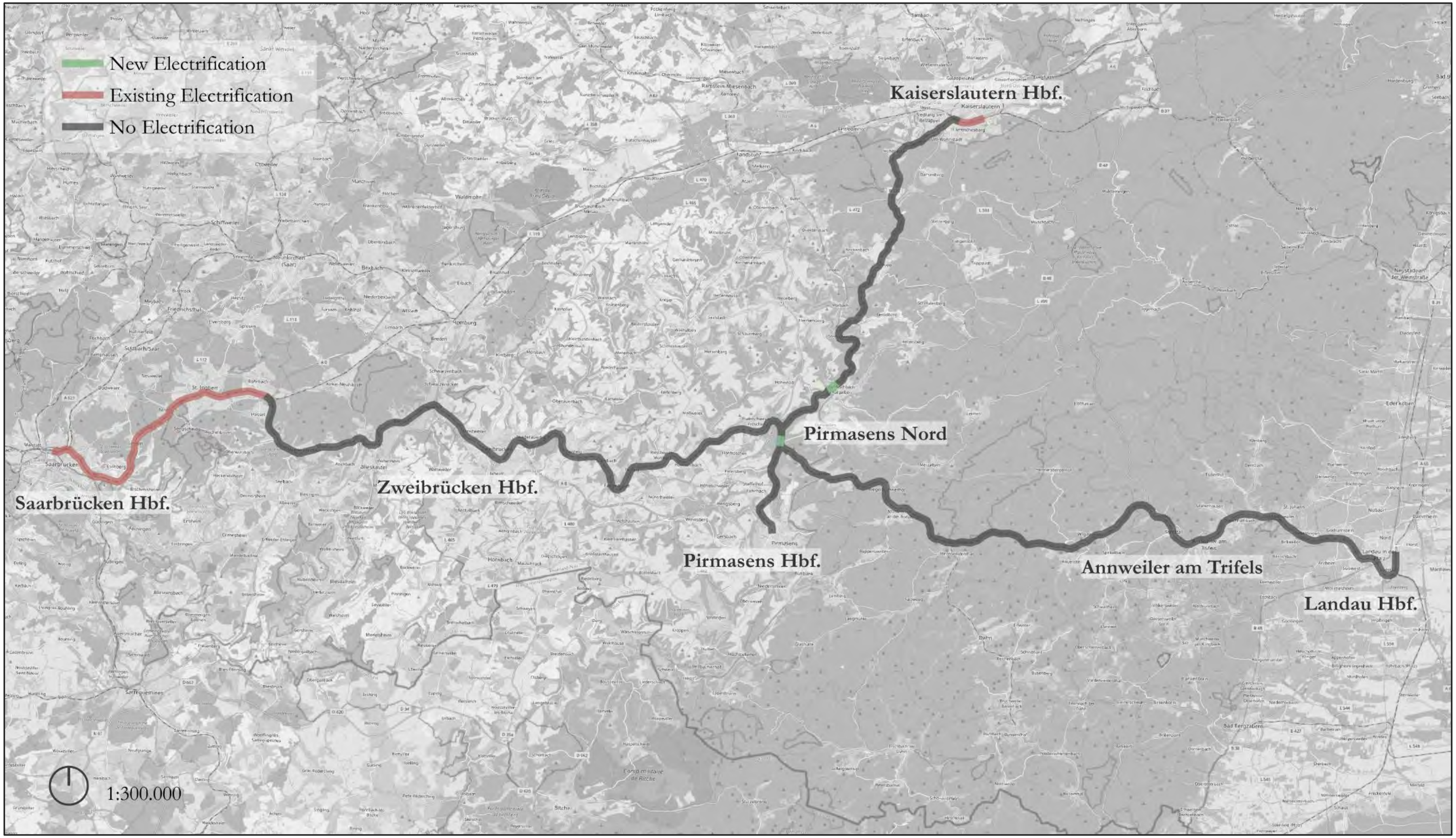

Figure 79: Existing and proposed electrification for model run “large battery 3”. The electrification is given per track even though this is not visible at map scale.

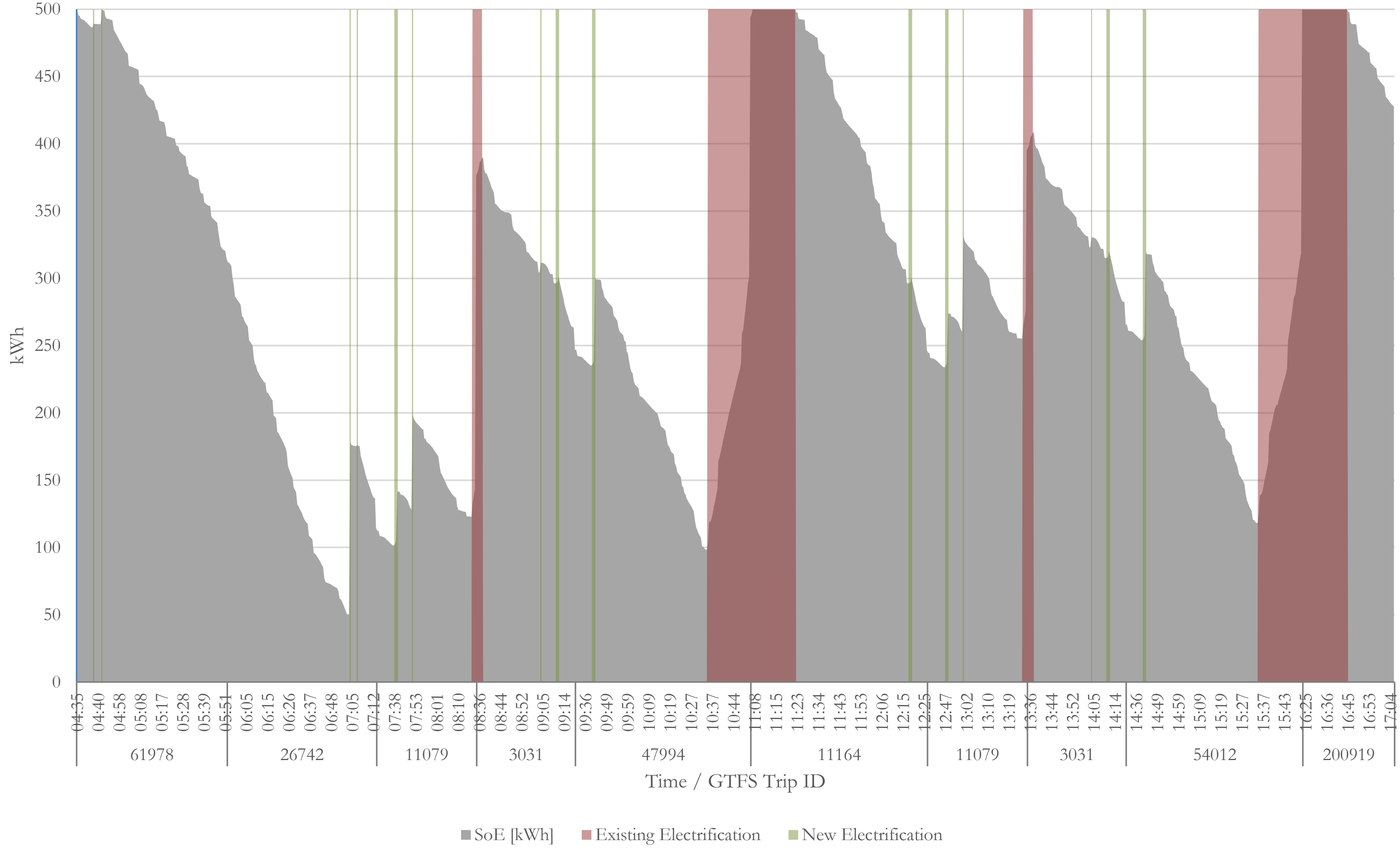

Figure 80: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run "large battery 3".

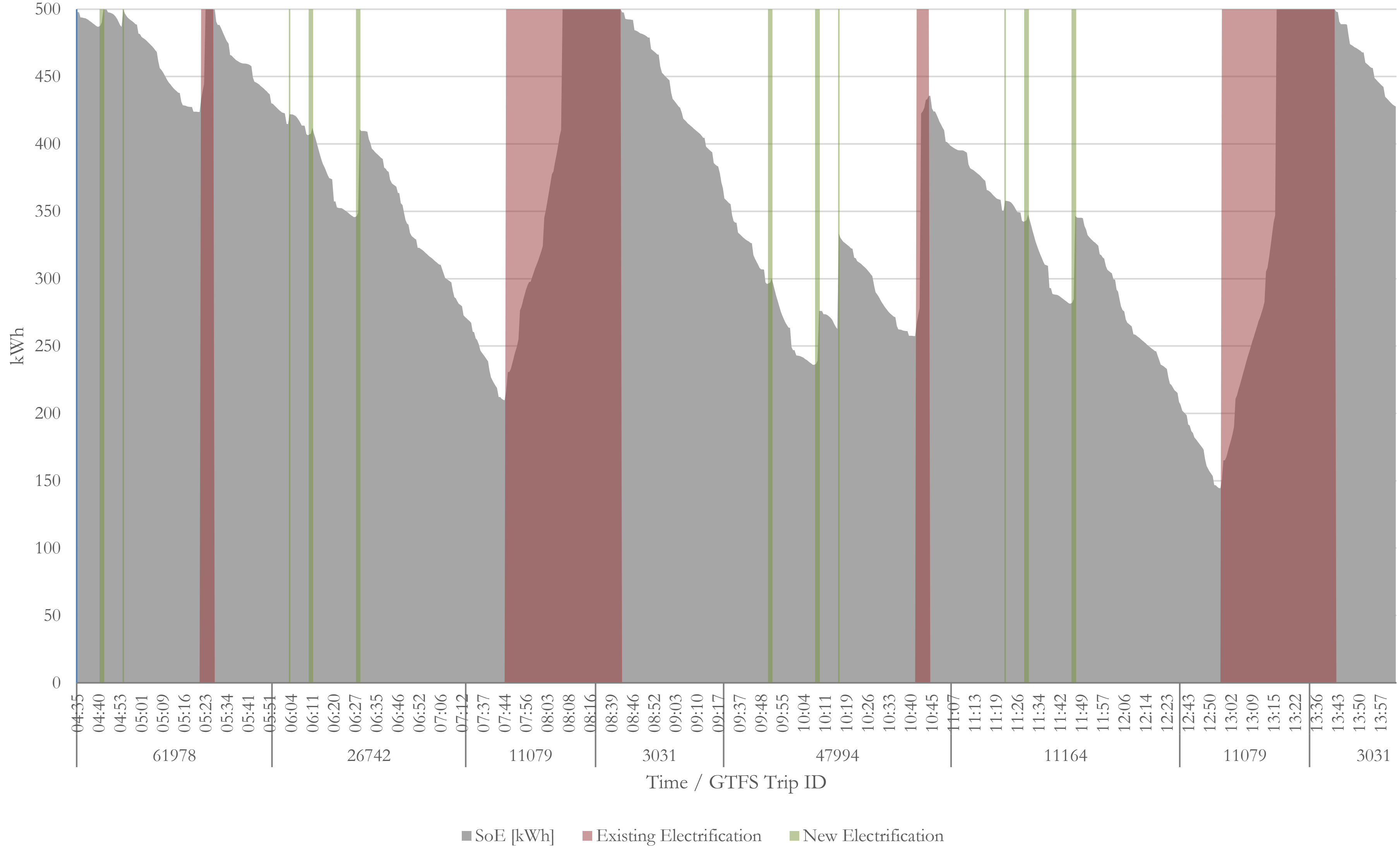

Figure 81: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run “large battery 3”.

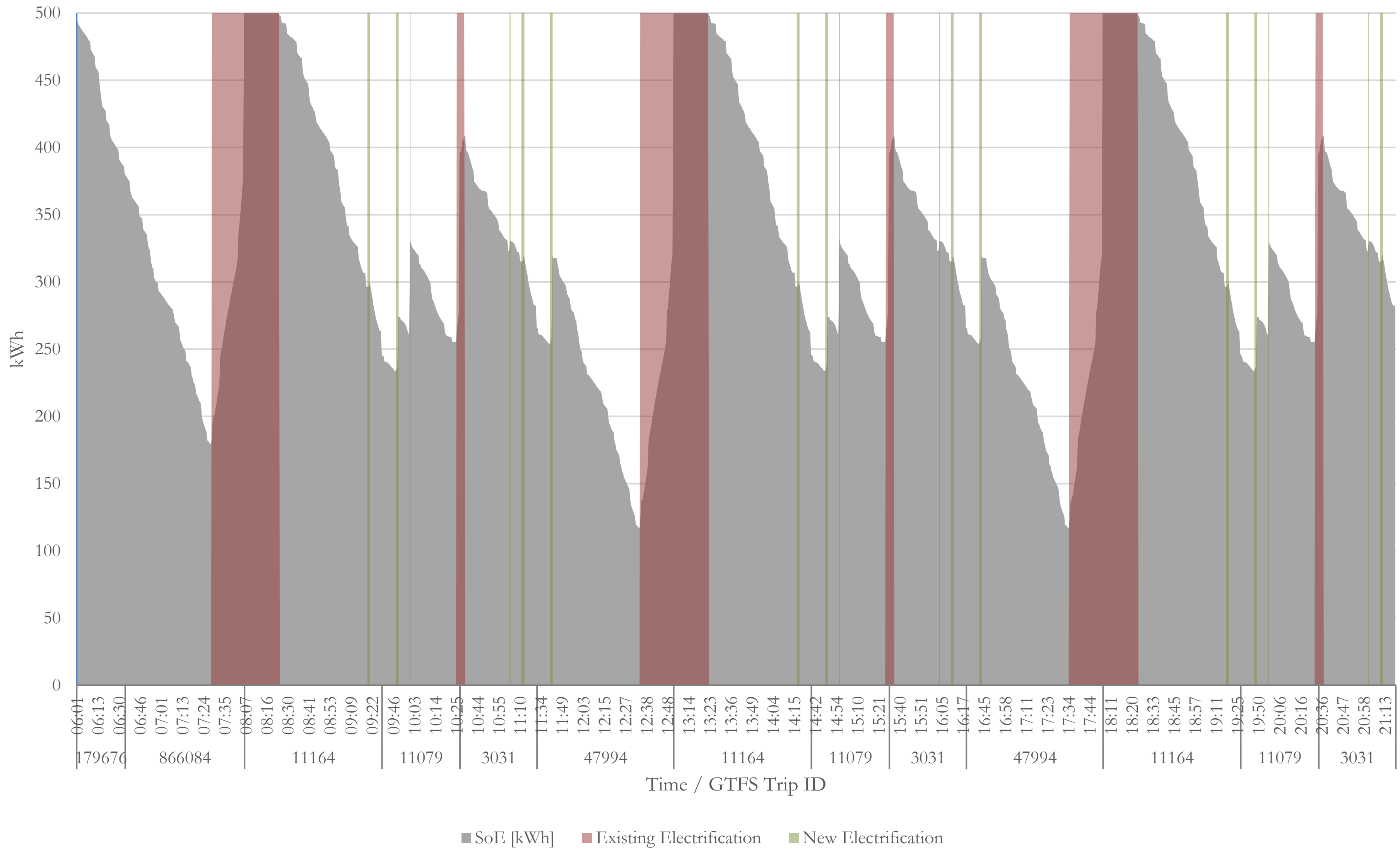

Figure 82: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run "large battery 3".

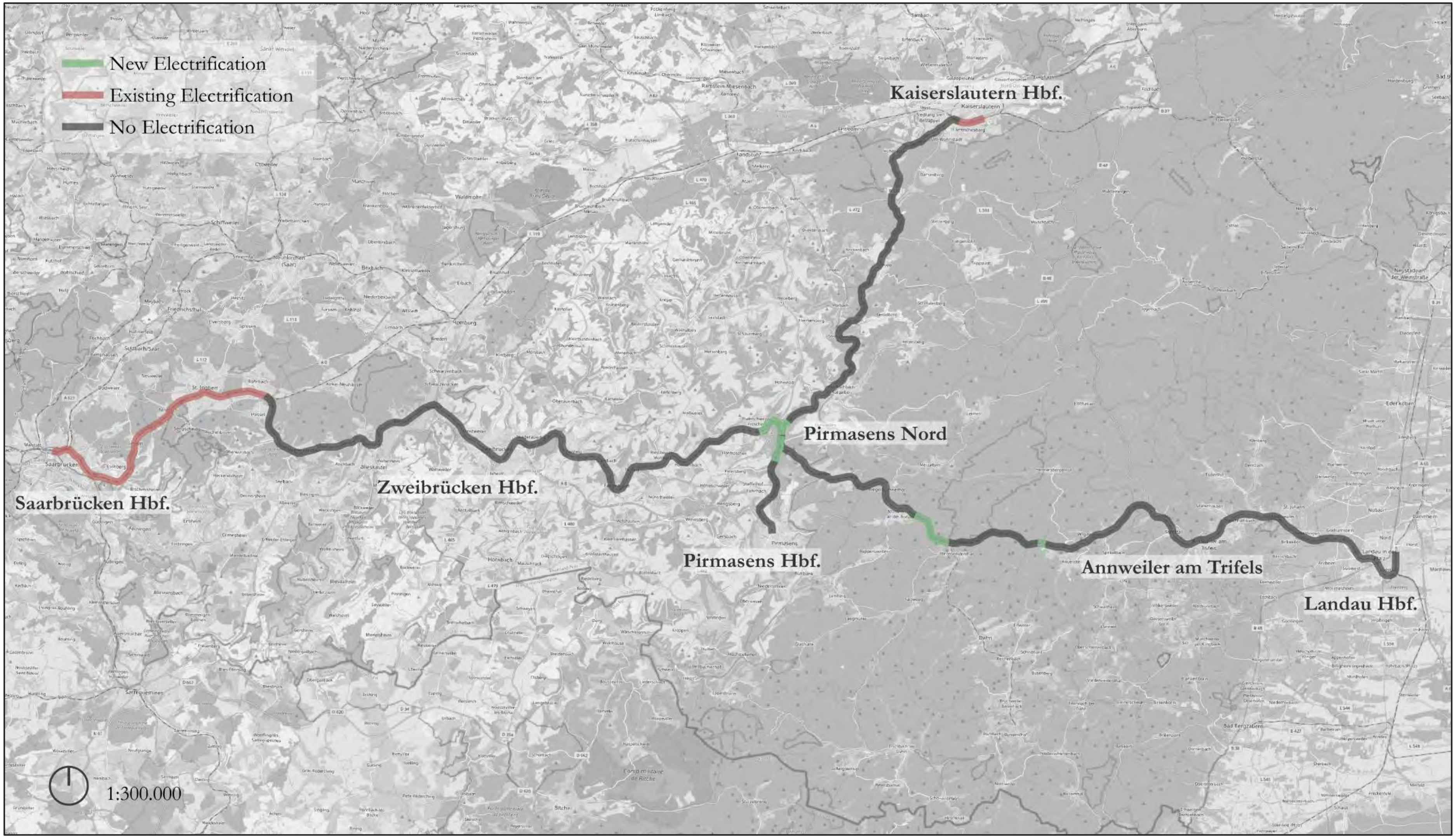

Figure 83: Existing and proposed electrification for model run "cheap substation 1". The electrification is given per track even though this is not visible at map scale.

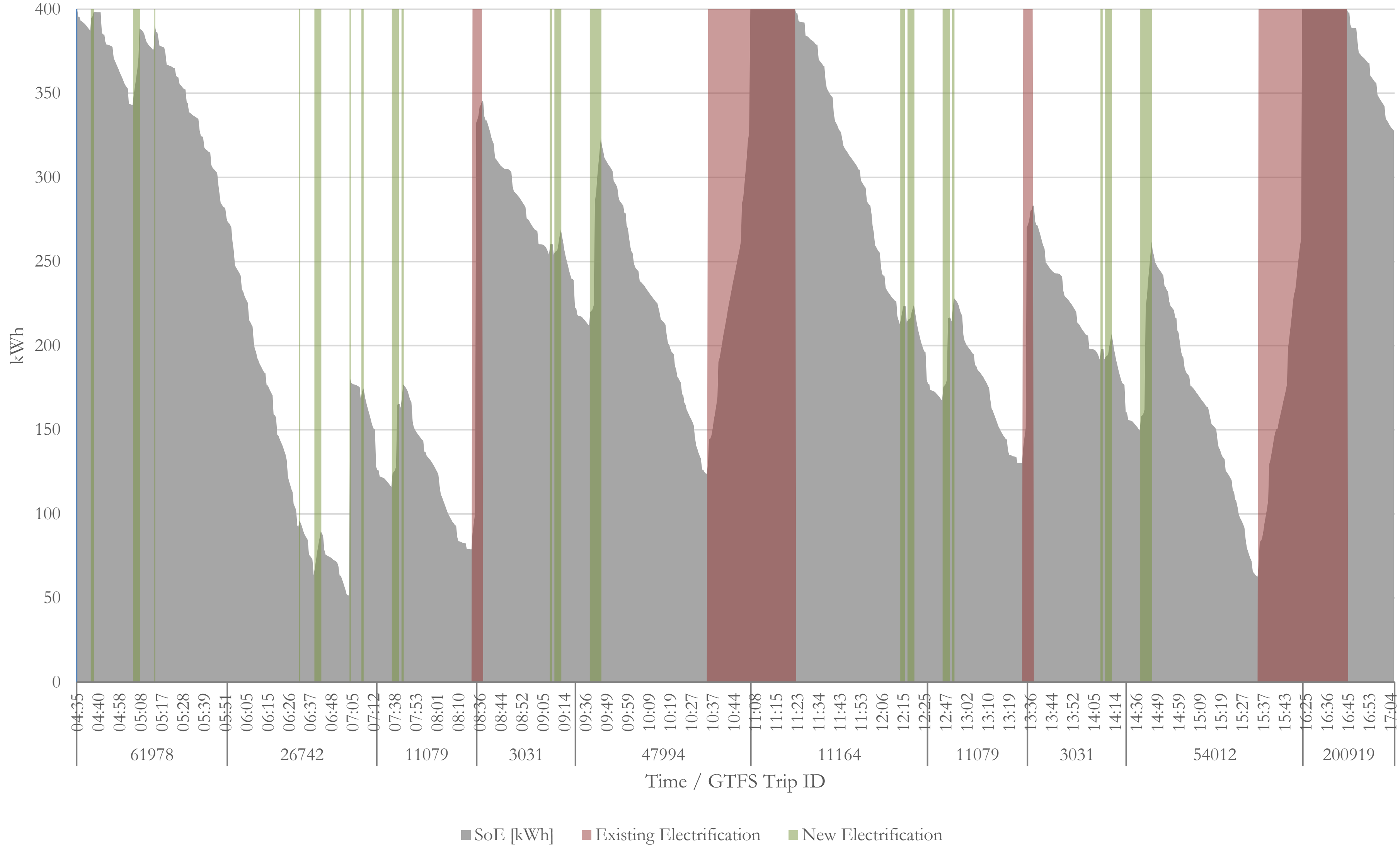

Figure 84: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run “cheap substation 1”.

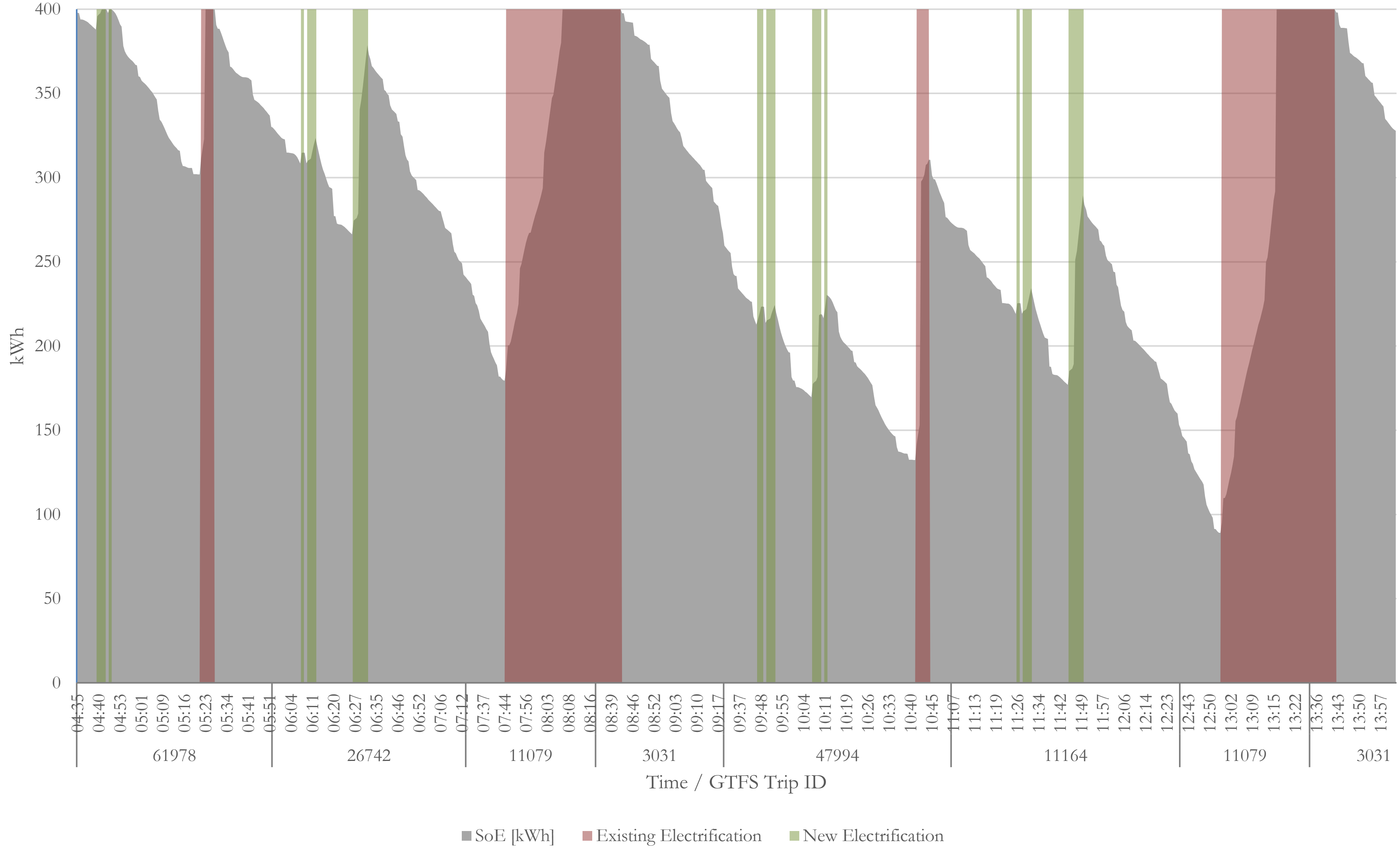

Figure 85: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run “cheap substation 1”.

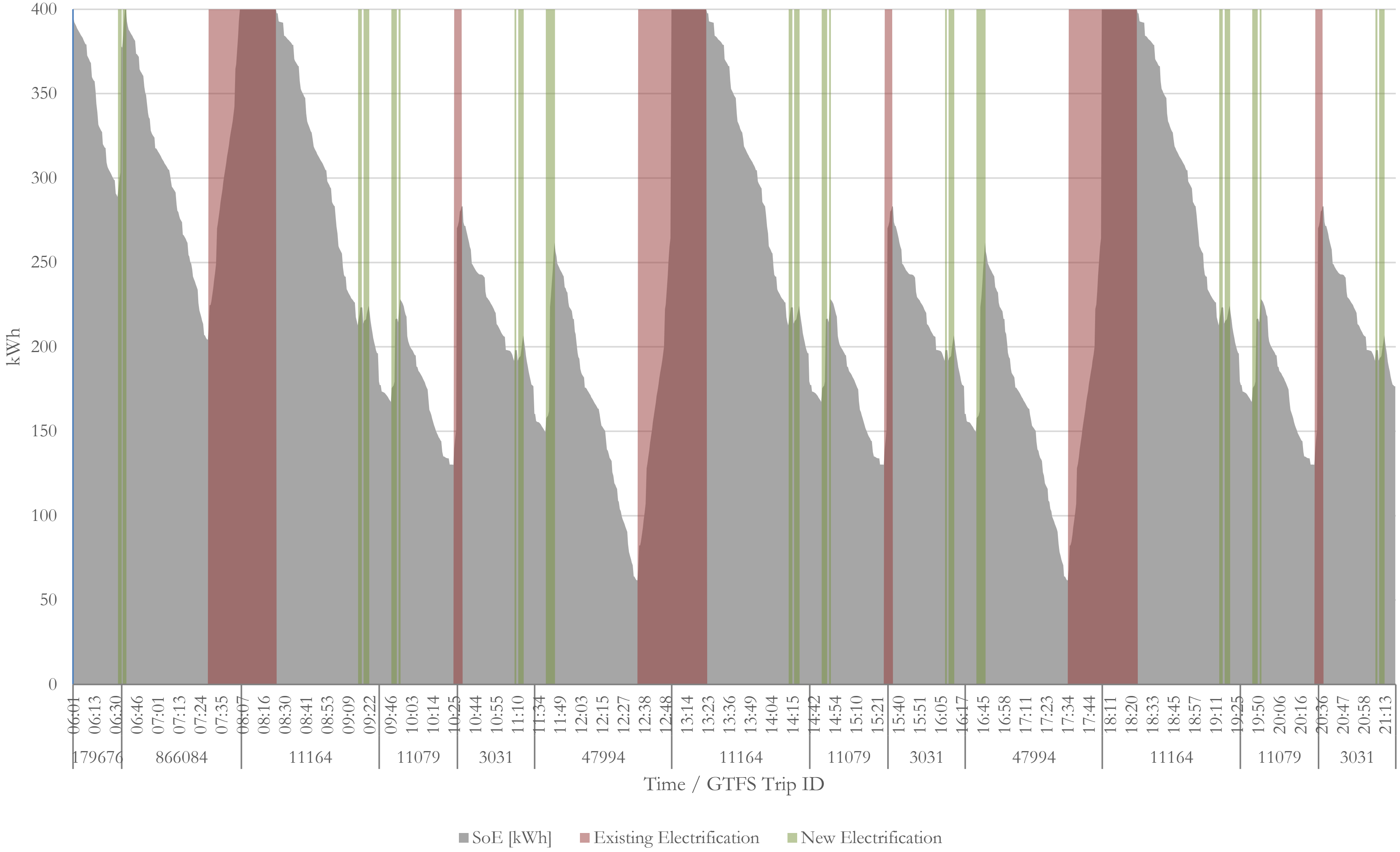

Figure 86: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run “cheap substation 1”.

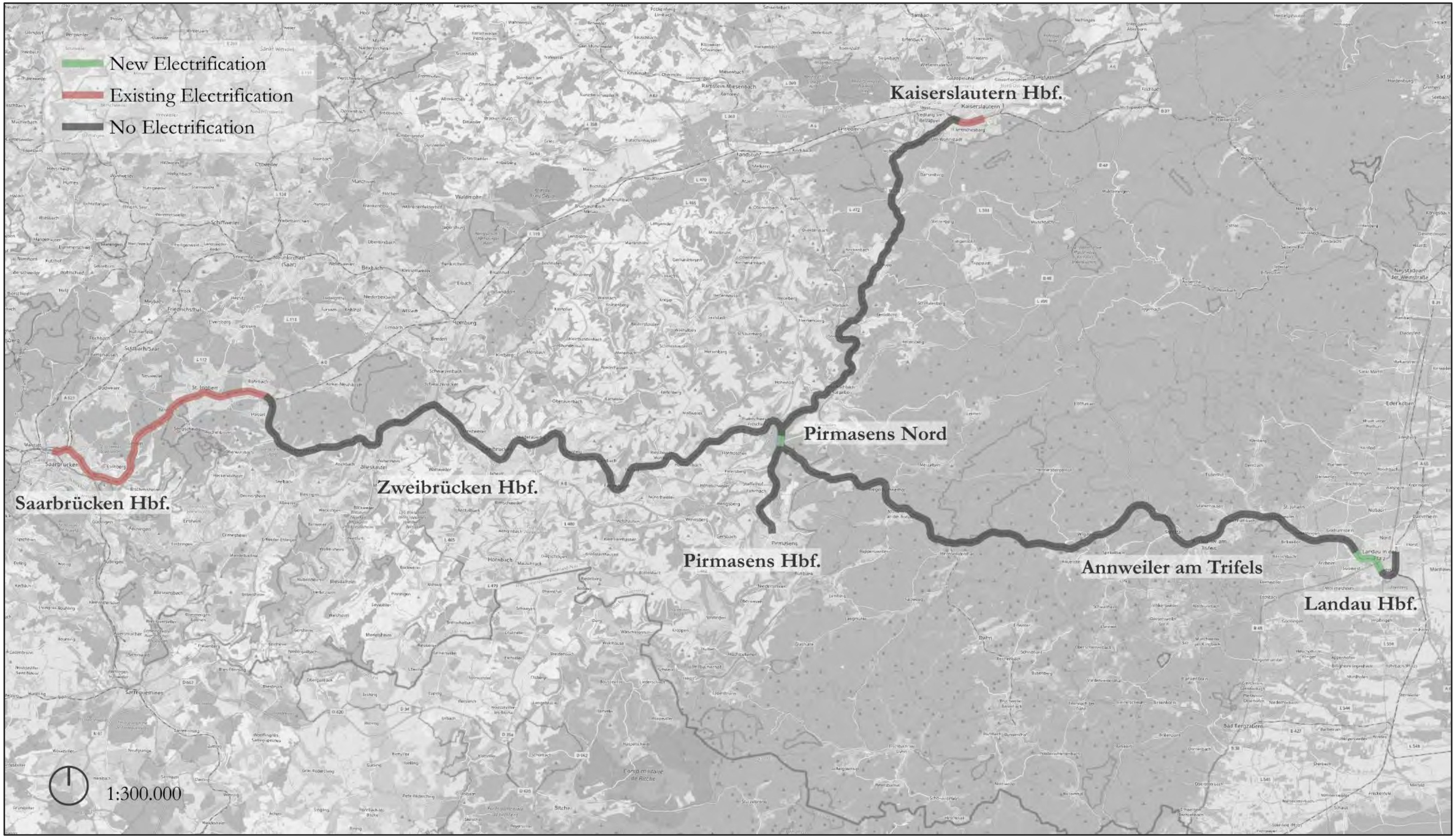

Figure 87: Existing and proposed electrification for model run “cheap substation 2”. The electrification is given per track even though this is not visible at map scale.

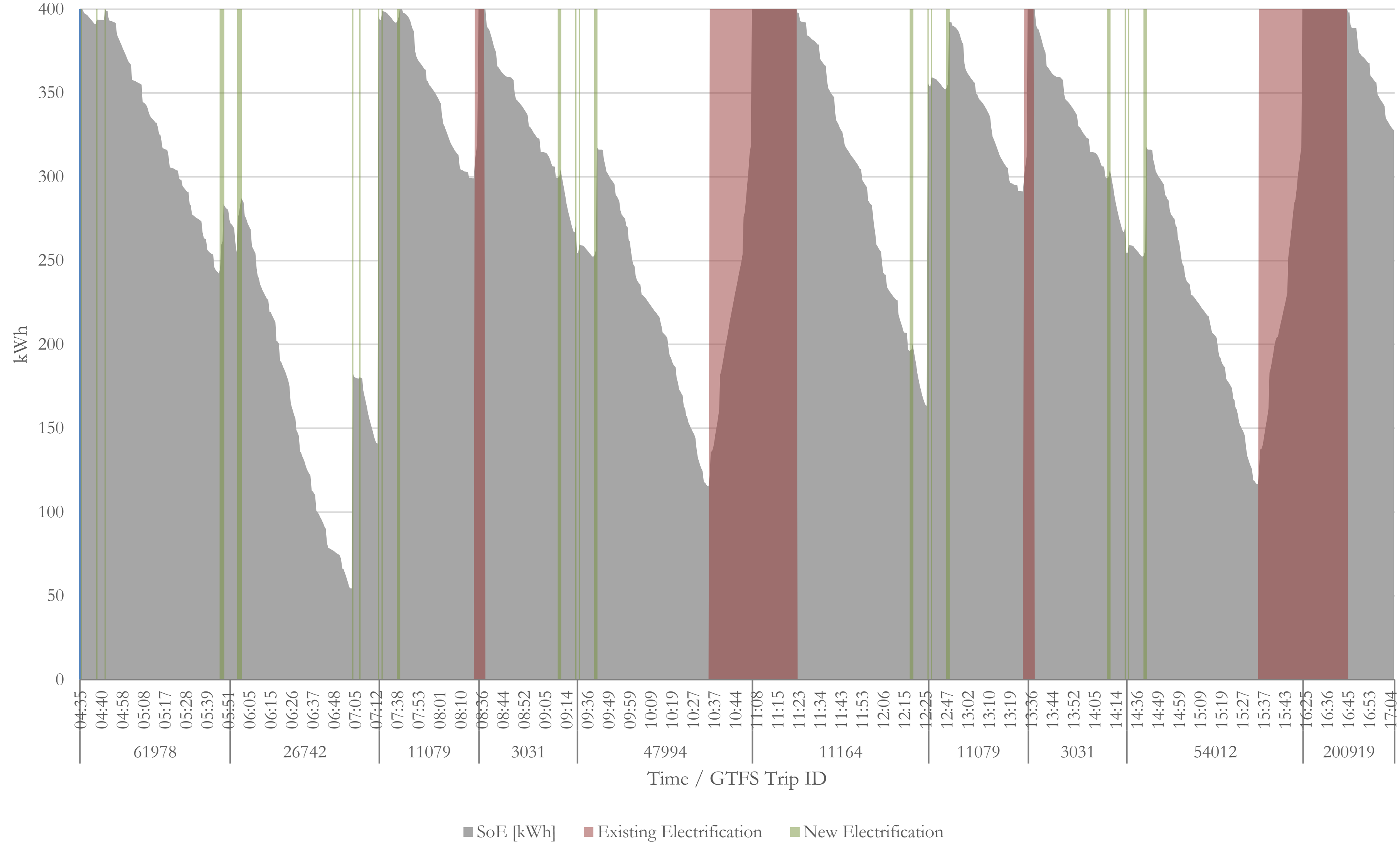

Figure 88: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run "cheap substation 2".

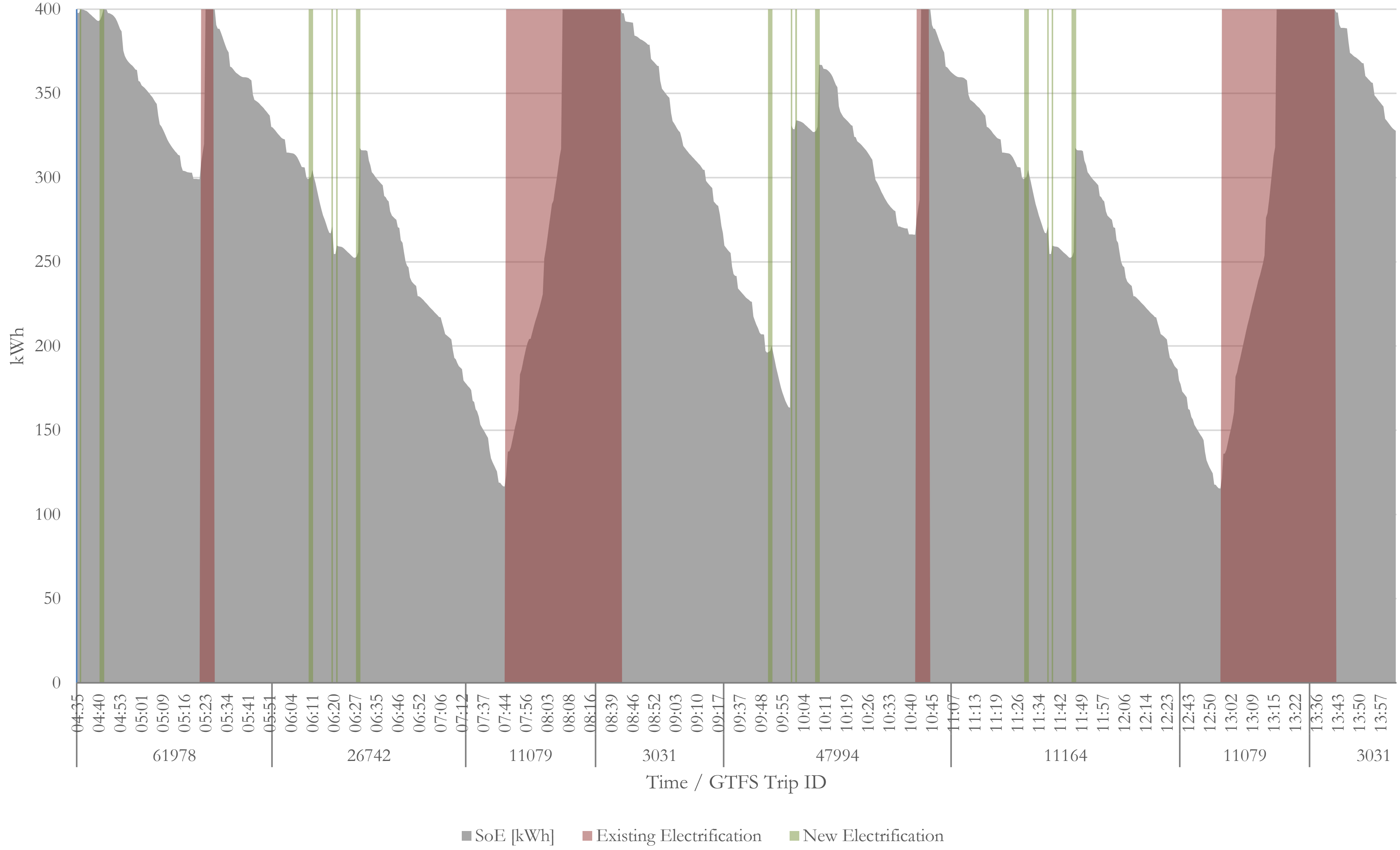

Figure 89: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run "cheap substation 2".

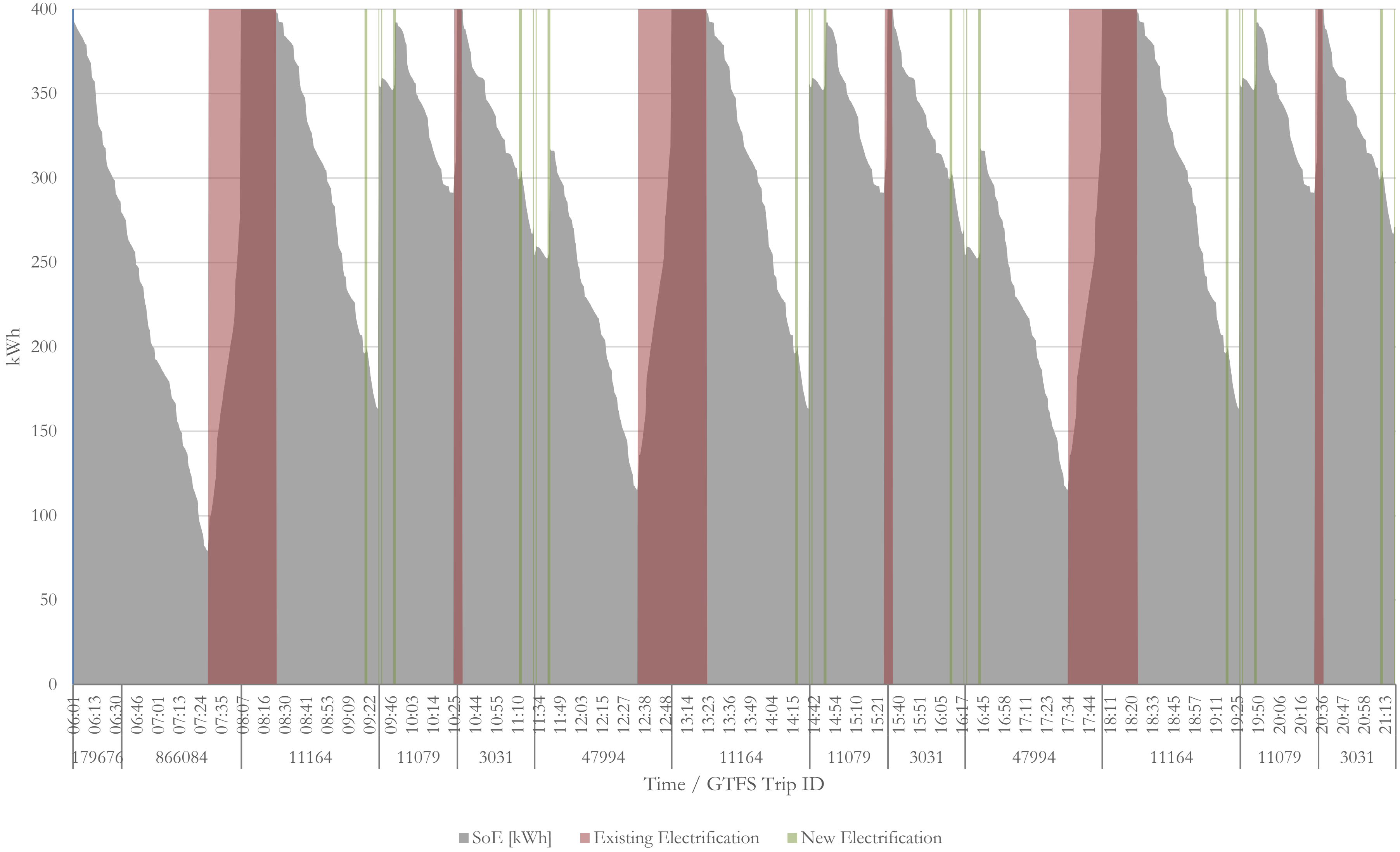

Figure 90: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run "cheap substation 2".

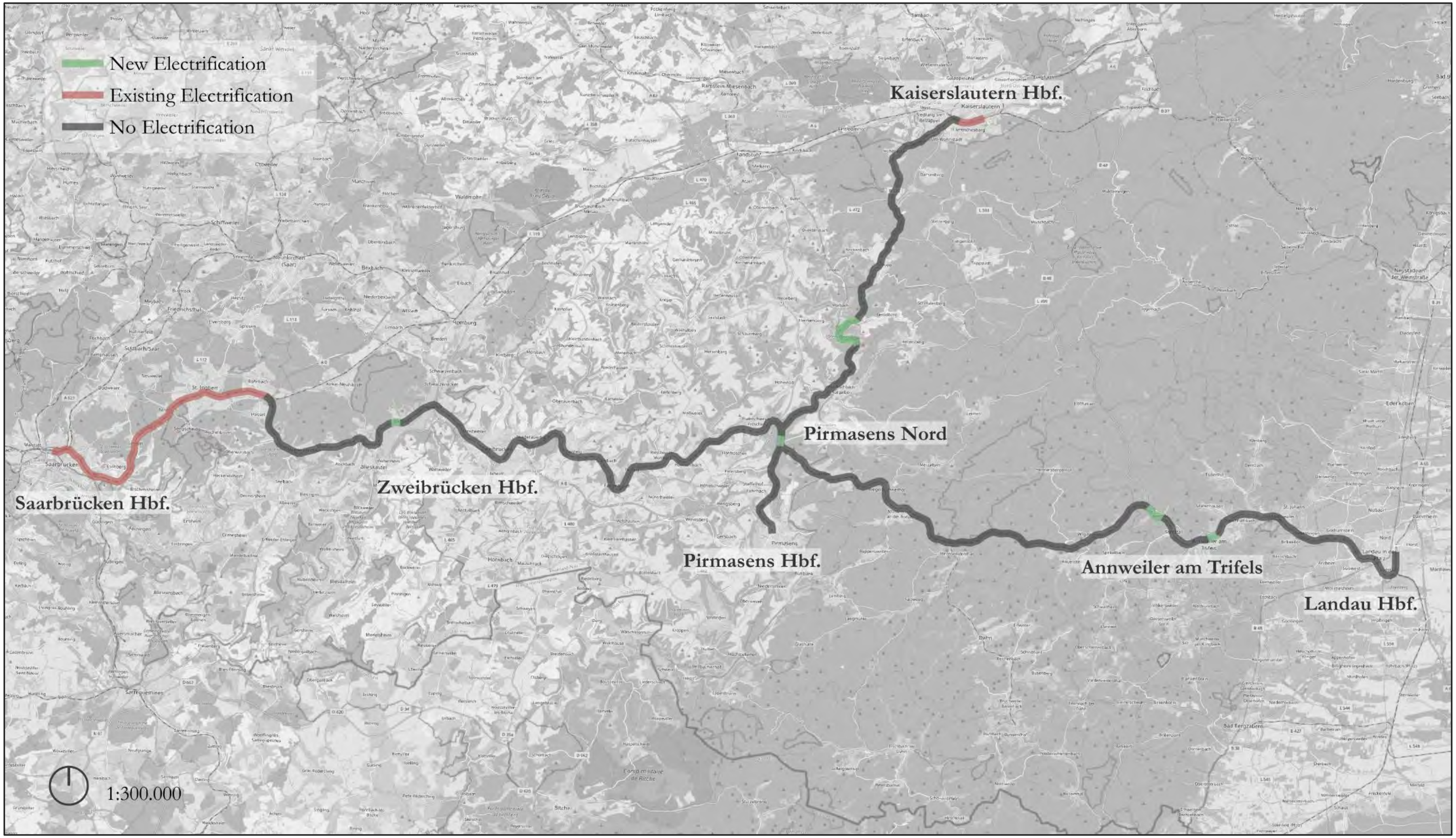

Figure 91: Existing and proposed electrification for model run "cheap substation 3". The electrification is given per track even though this is not visible at map scale.

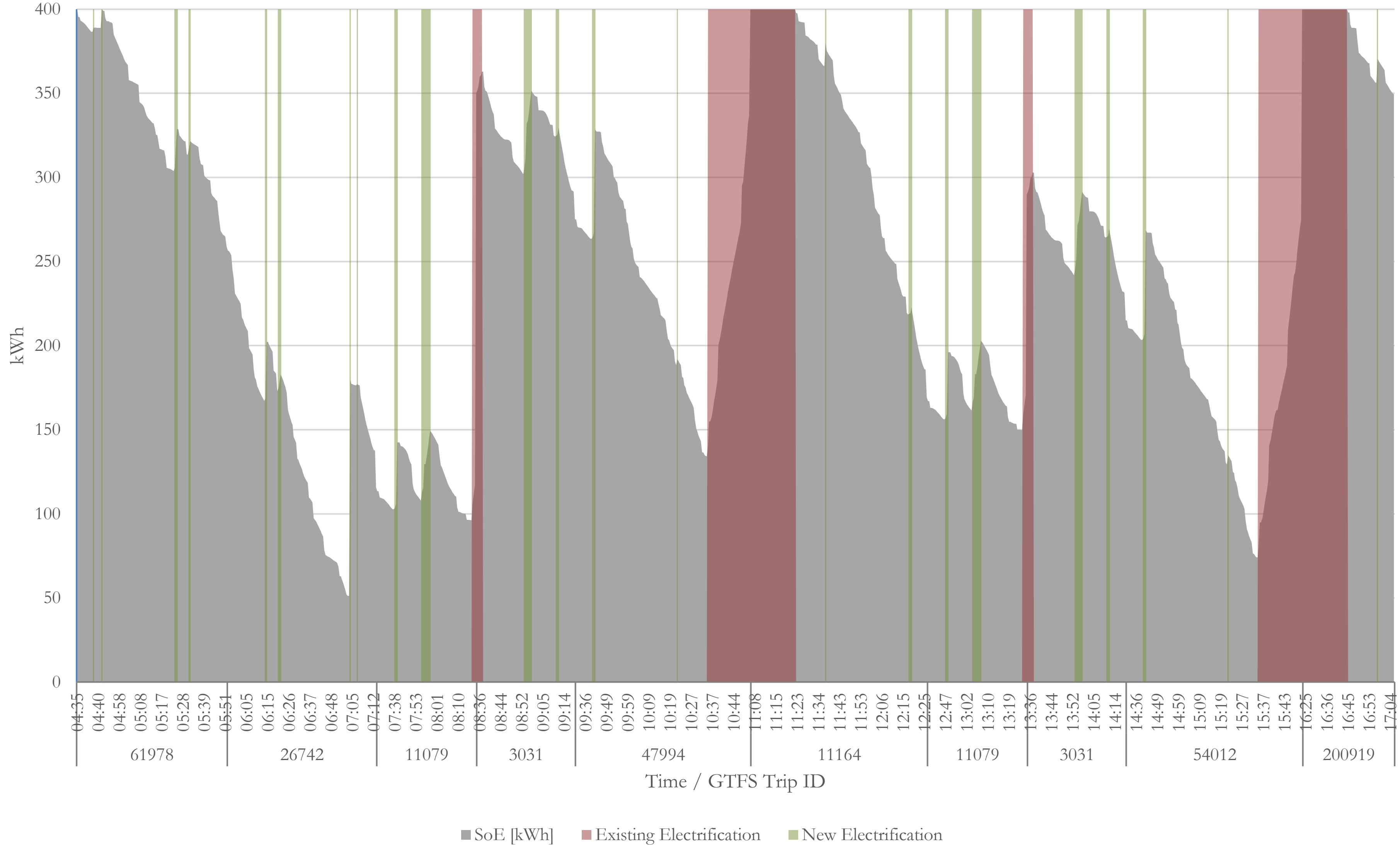

Figure 92: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run “cheap substation 3”.

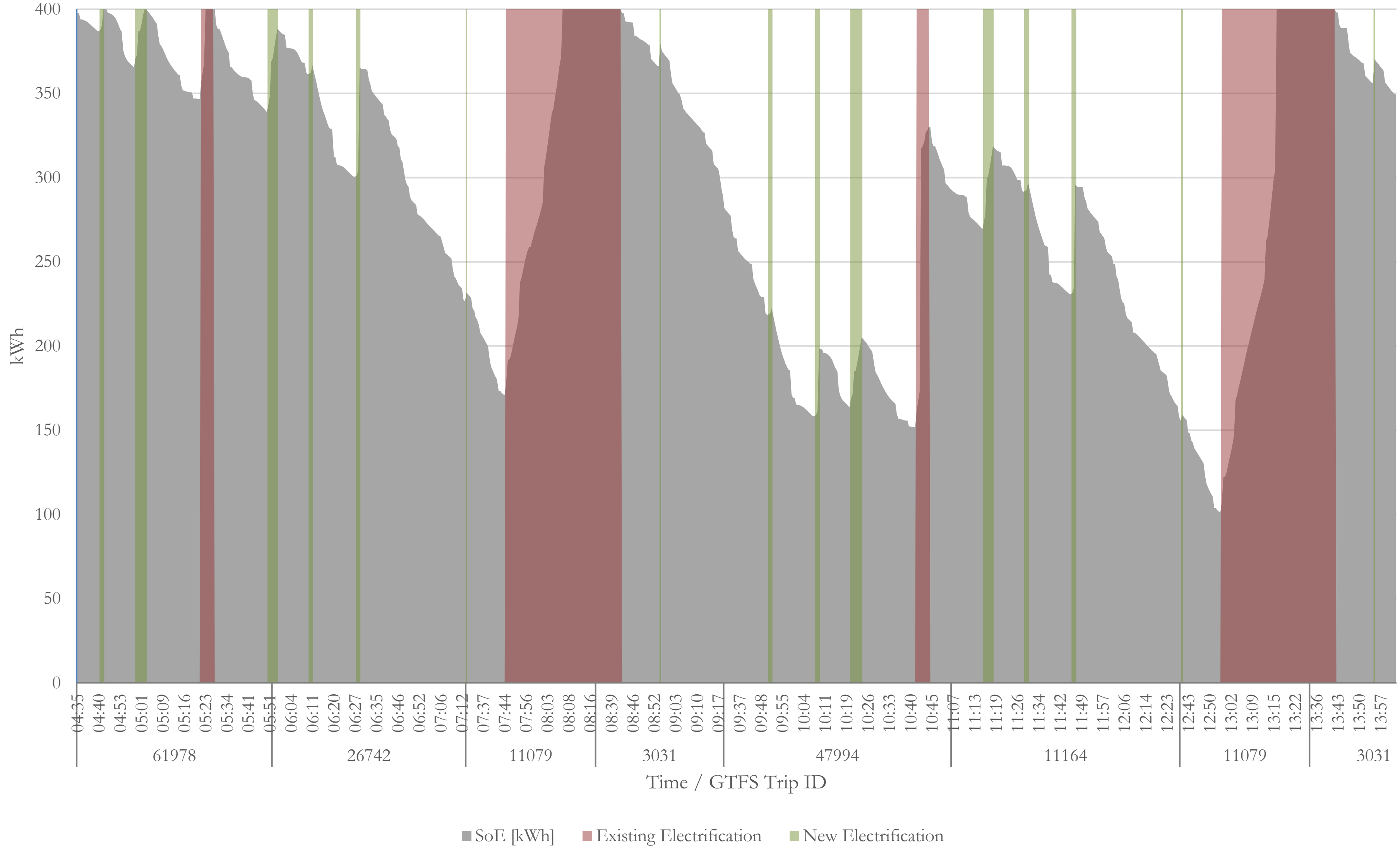

Figure 93: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run “cheap substation 3”.

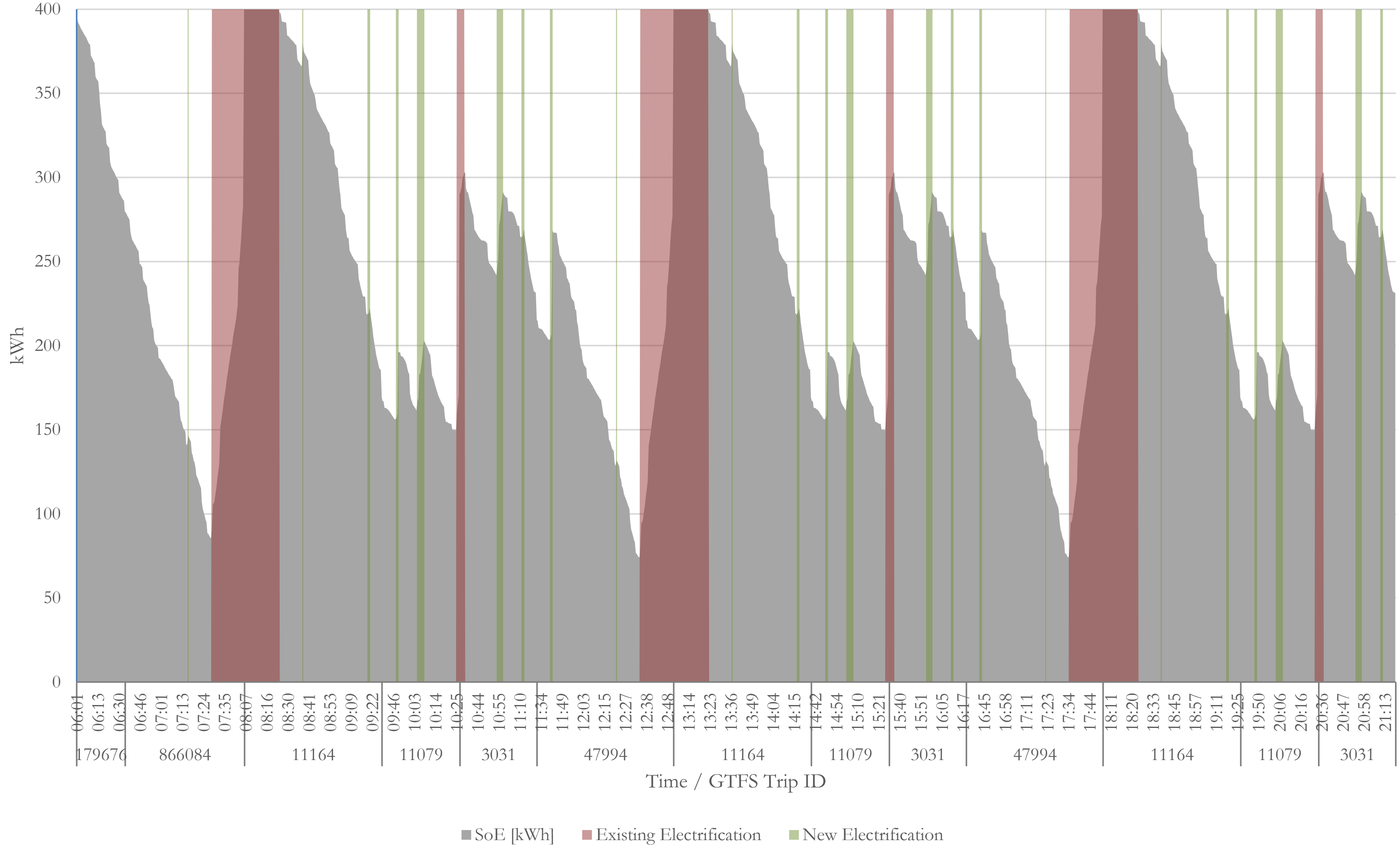

Figure 94: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run "cheap substation 3".

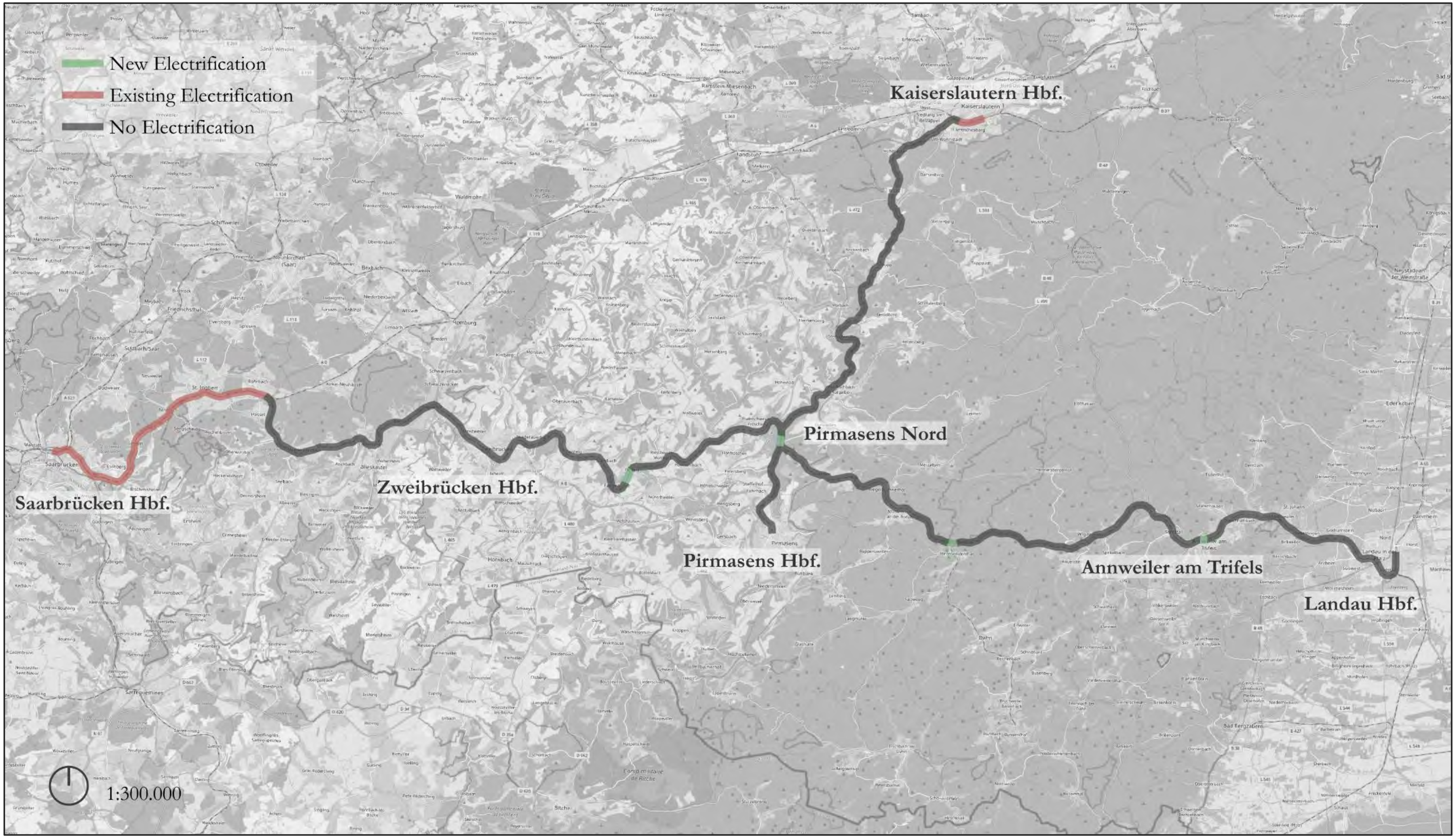

Figure 95: Existing and proposed electrification for model run "standard substation 1". The electrification is given per track even though this is not visible at map scale.

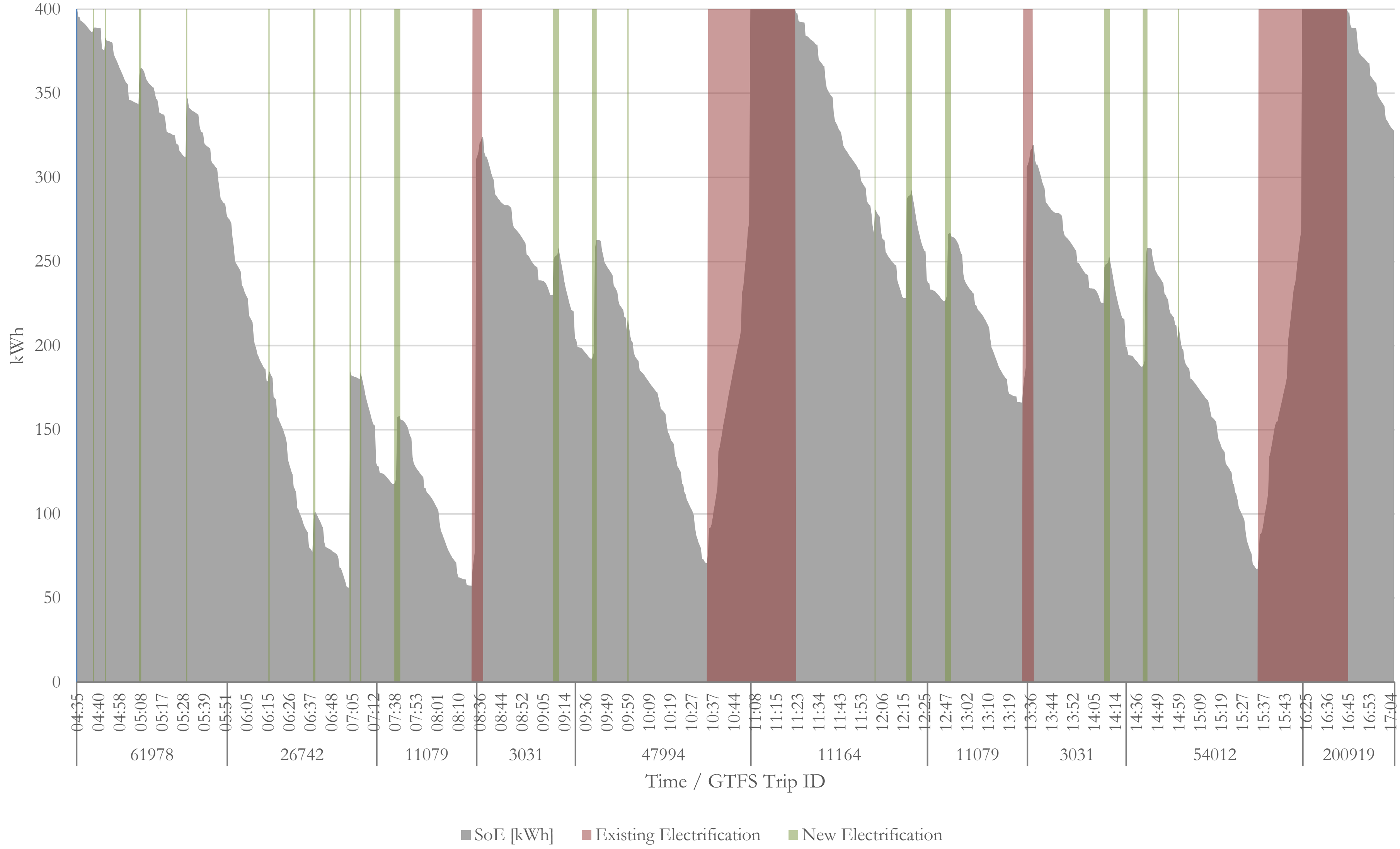

Figure 96: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run "standard substation 1".

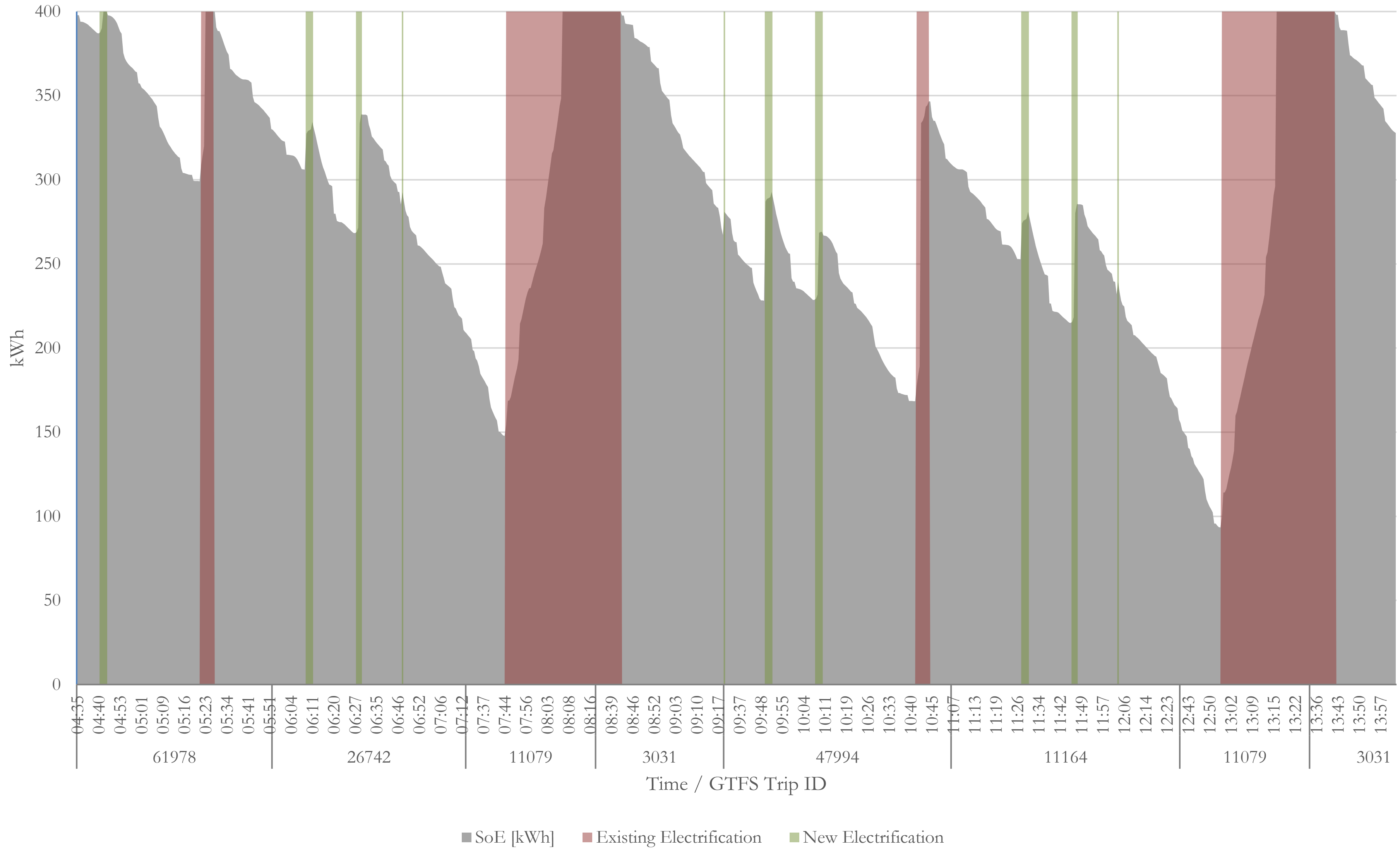

Figure 97: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run “standard substation 1”.

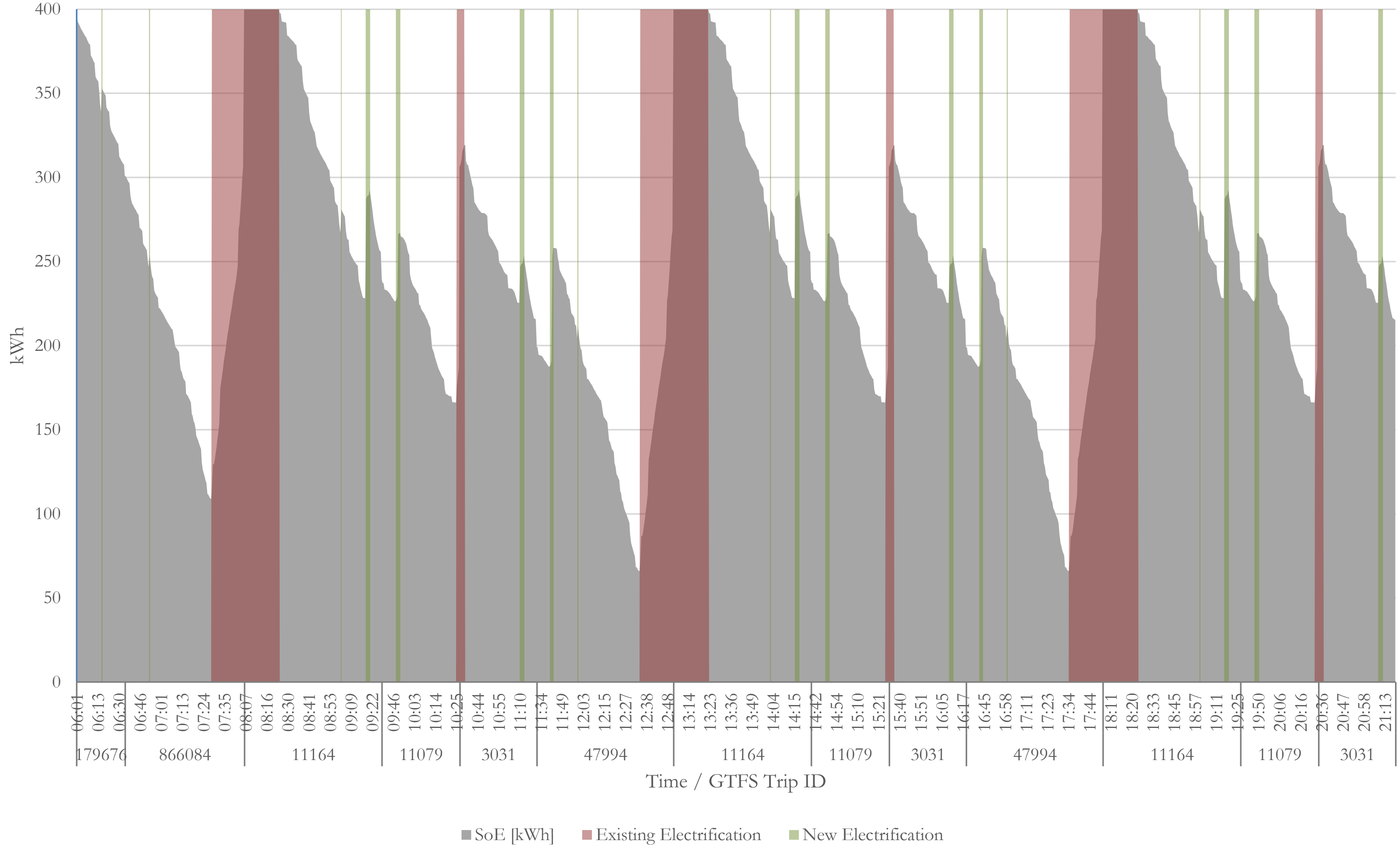

Figure 98: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run "standard substation 1".

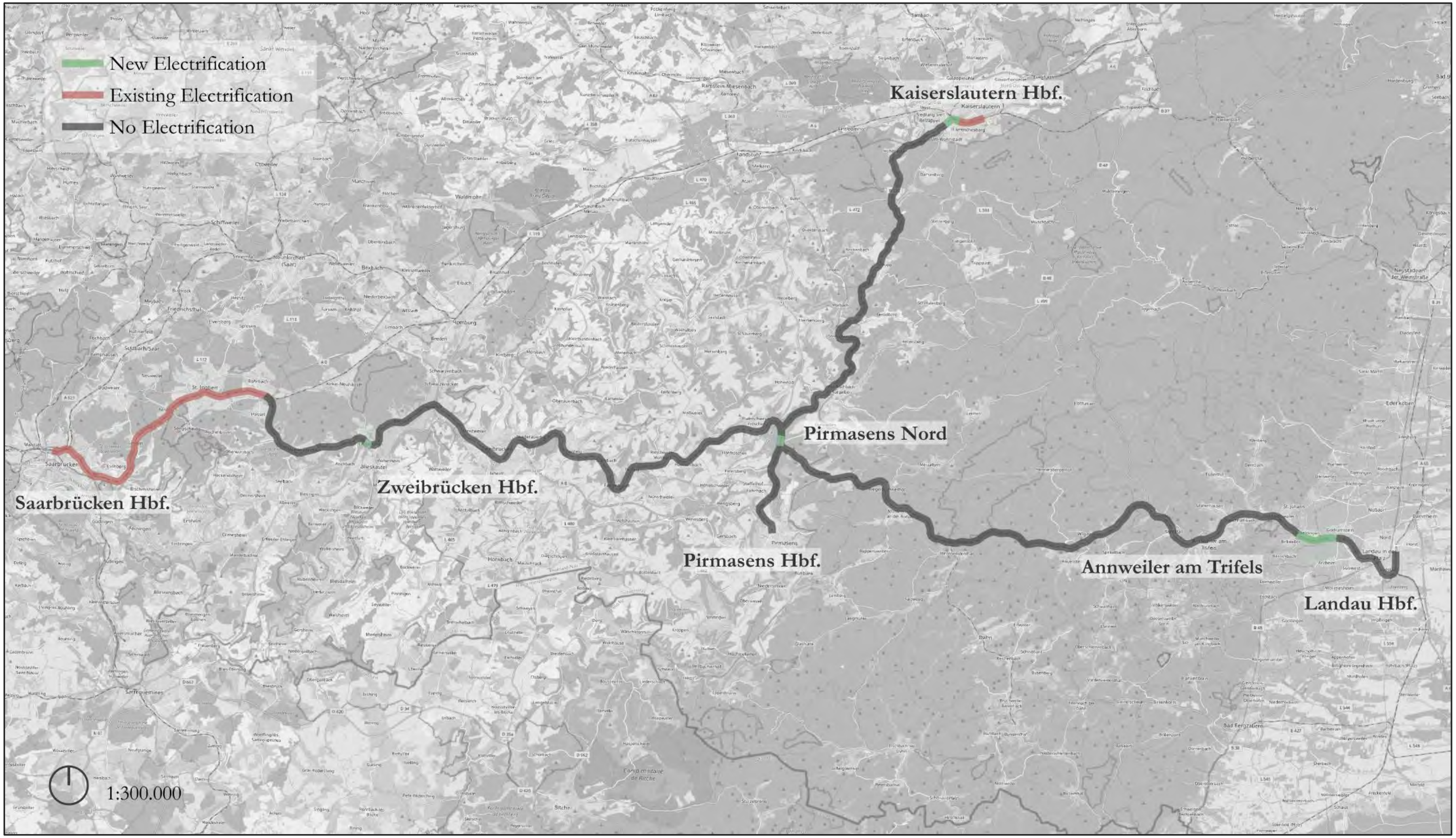

Figure 99: Existing and proposed electrification for model run "standard substation 2". The electrification is given per track even though this is not visible at map scale.

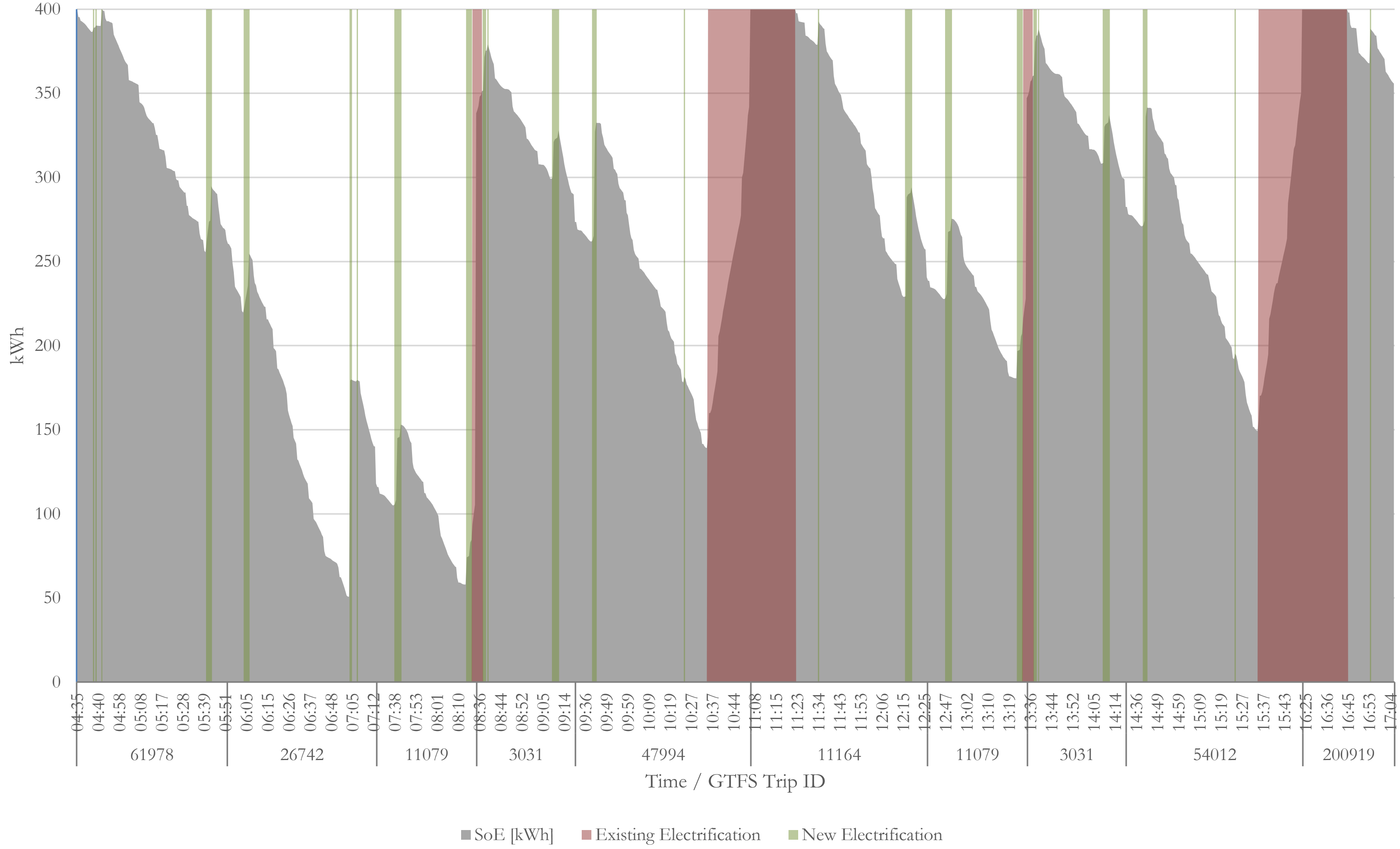

Figure 100: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run "standard substation 2".

Figure 101: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run “standard substation 2”.

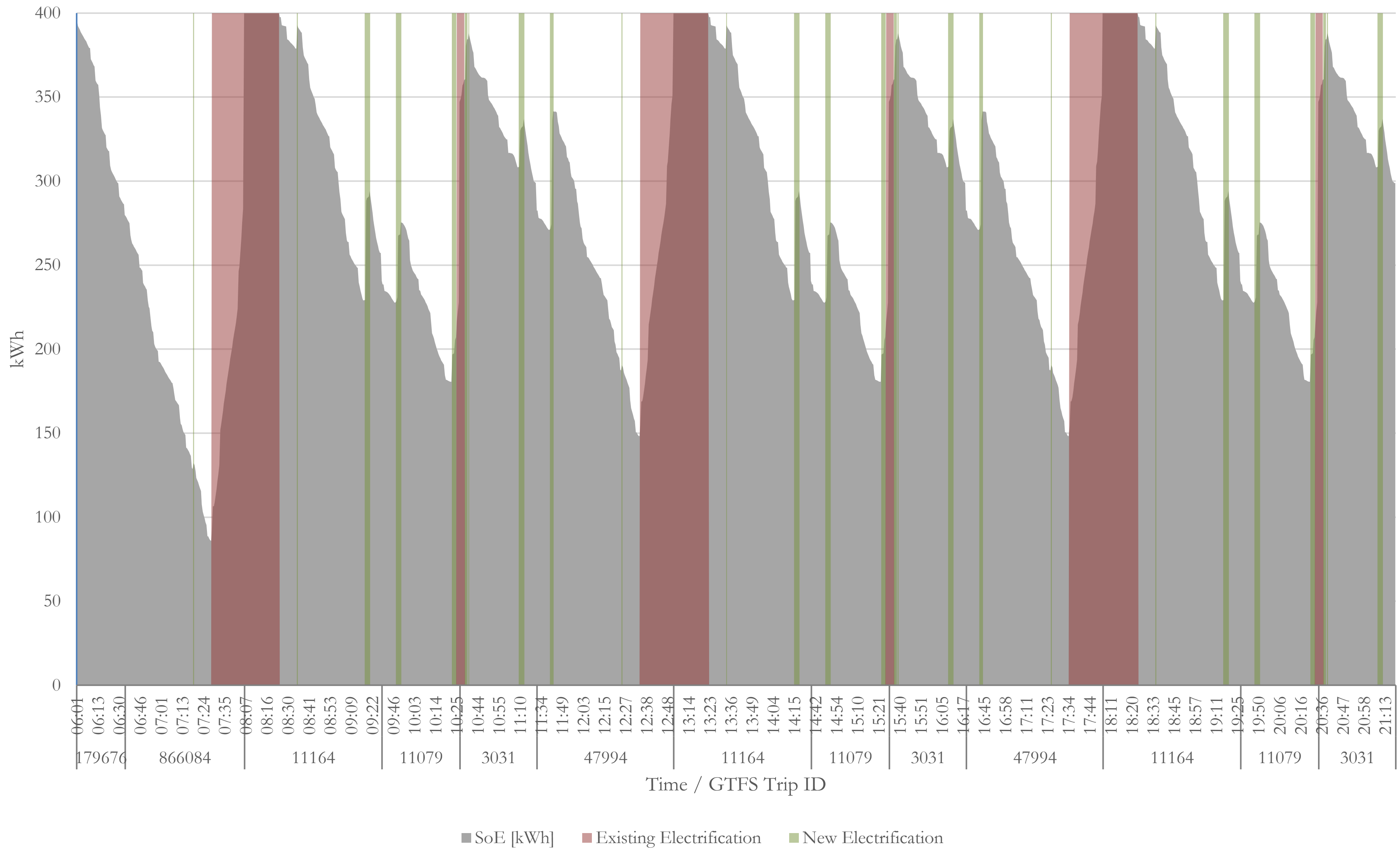

Figure 102: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run “standard substation 2”.

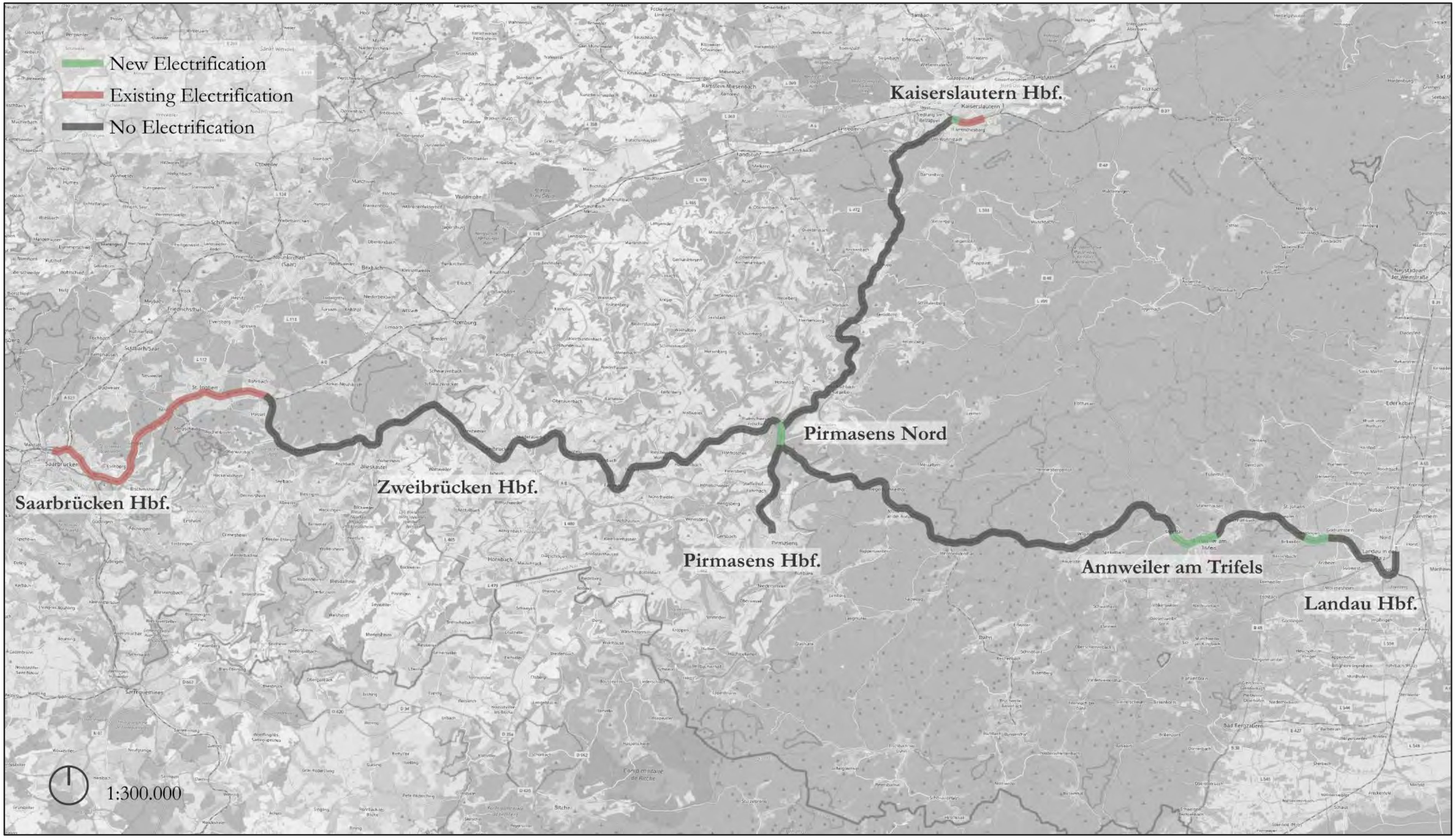

Figure 103: Existing and proposed electrification for model run “standard substation 3”. The electrification is given per track even though this is not visible at map scale.

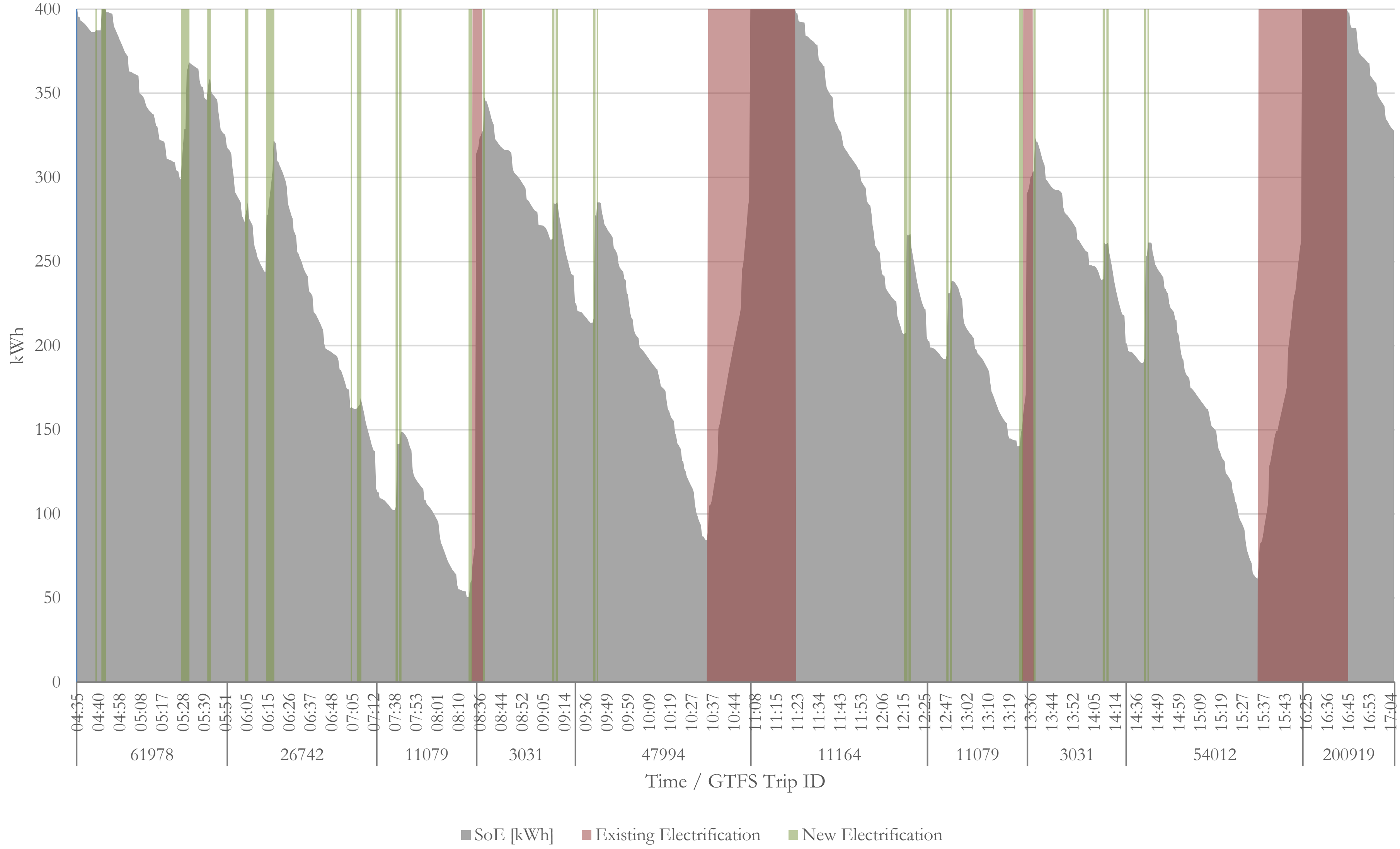

Figure 104: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run “standard substation 3”.

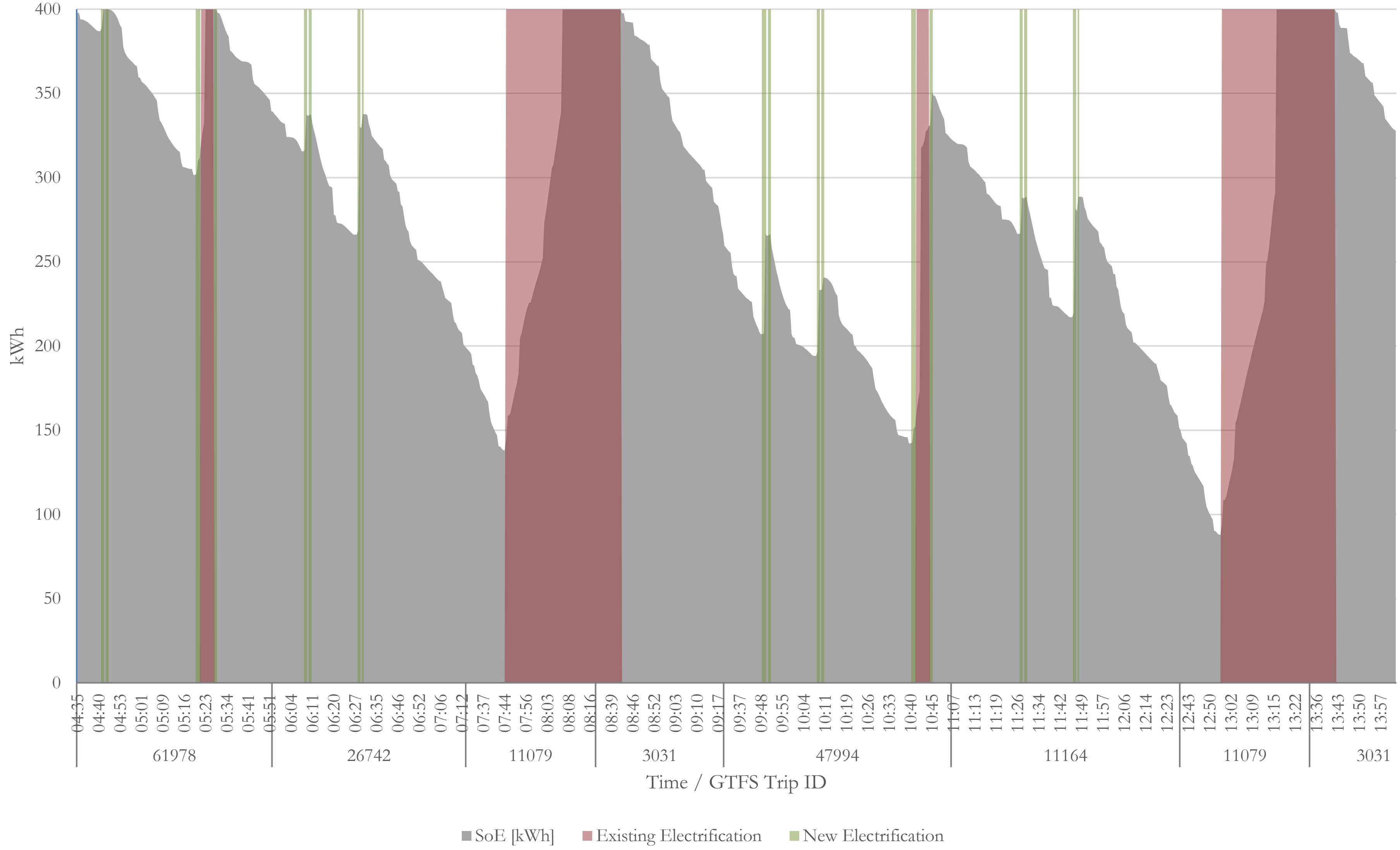

Figure 105: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run "standard substation 3".

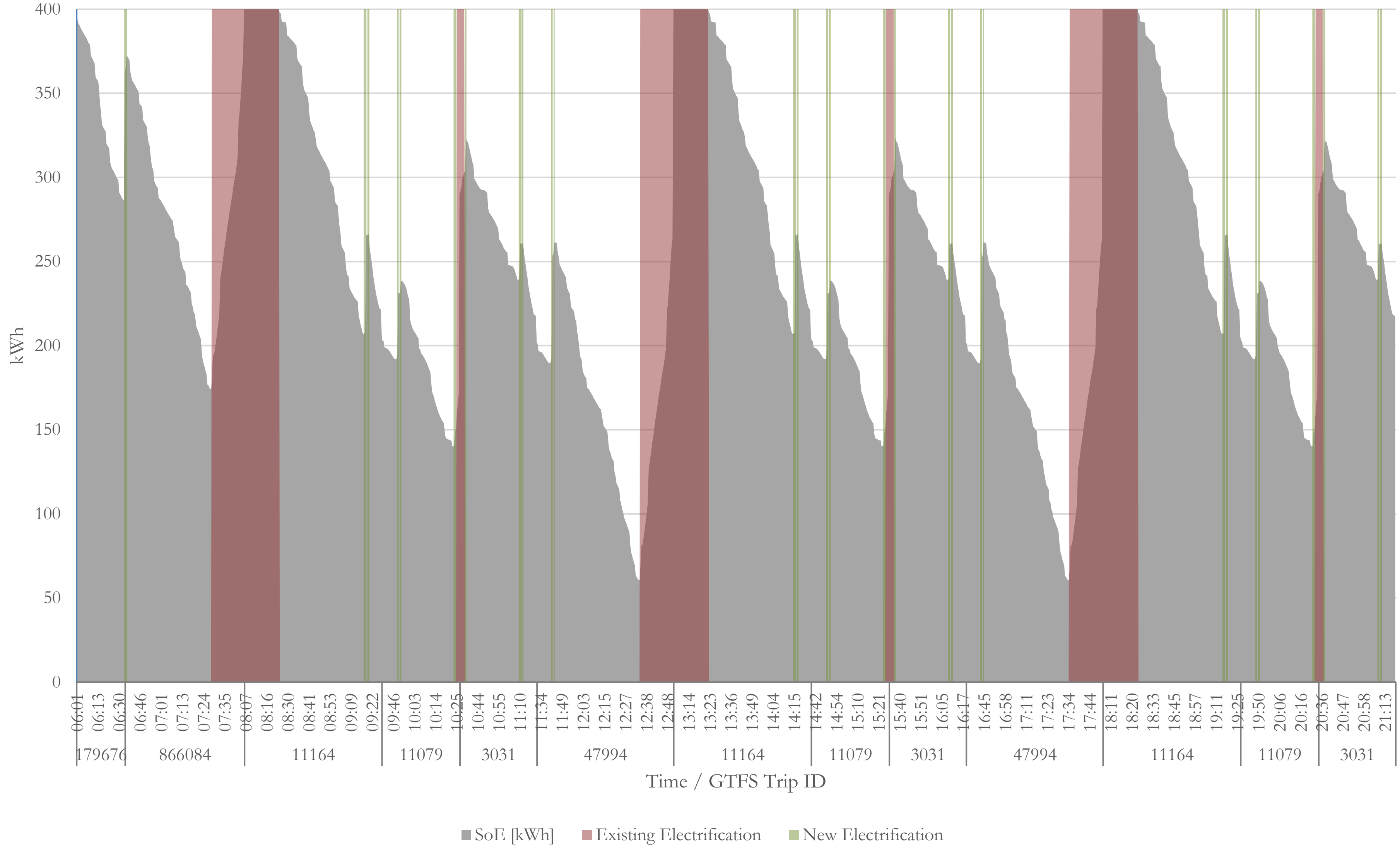

Figure 106: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run "standard substation 3".

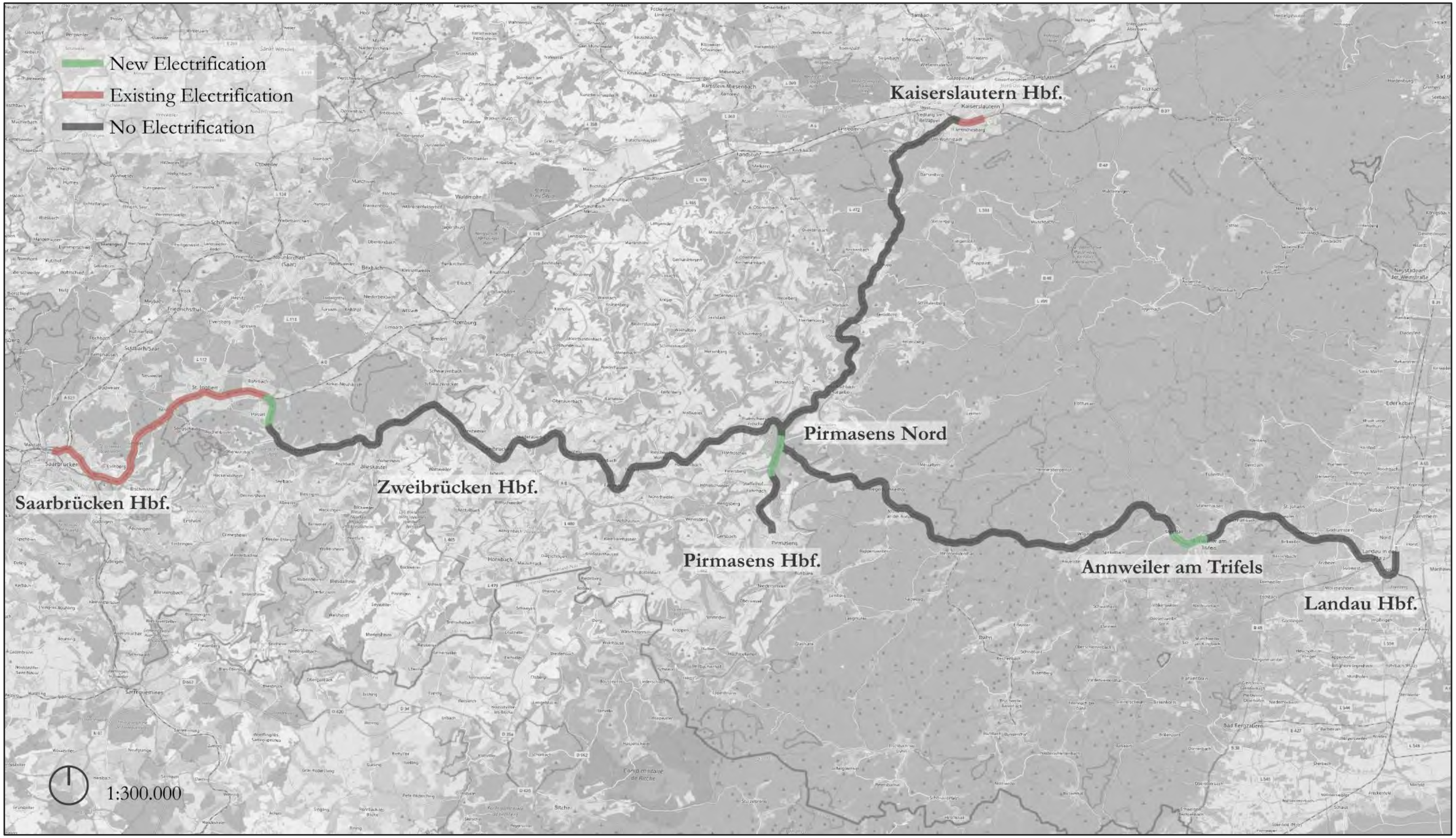

Figure 107: Existing and proposed electrification for model run "expensive substation 1". The electrification is given per track even though this is not visible at map scale.

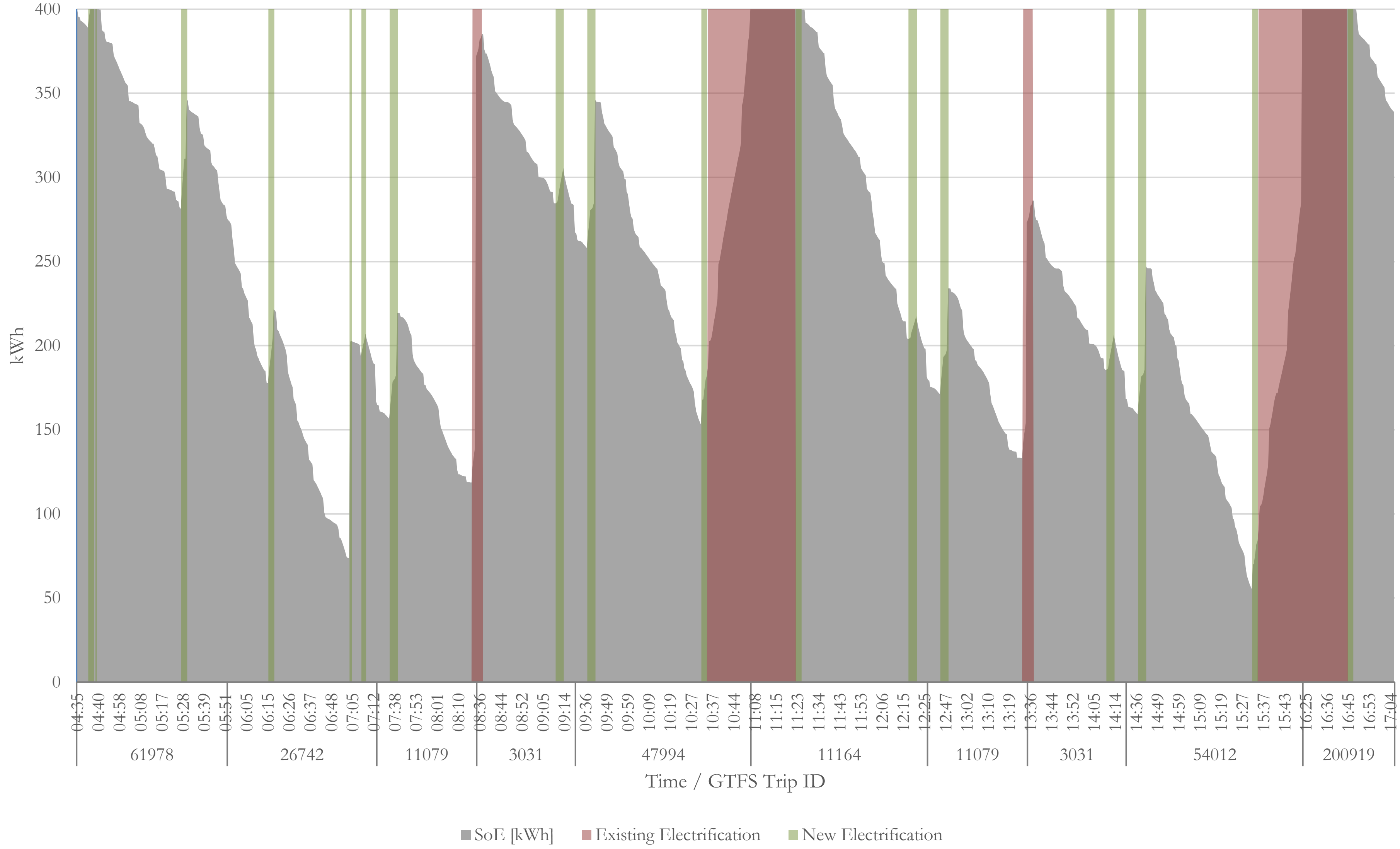

Figure 108: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run "expensive substation 1".

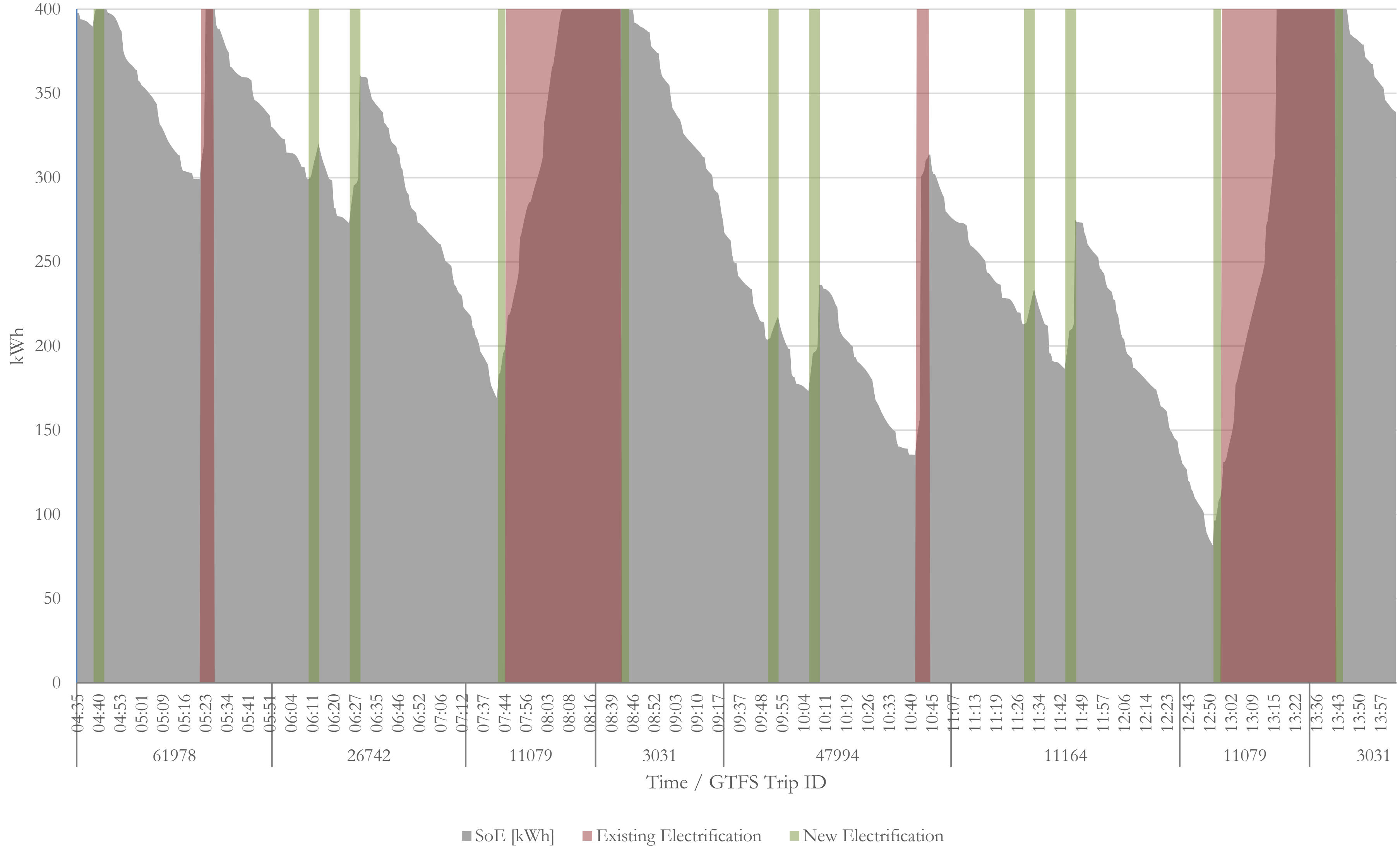

Figure 109: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run "expensive substation 1".

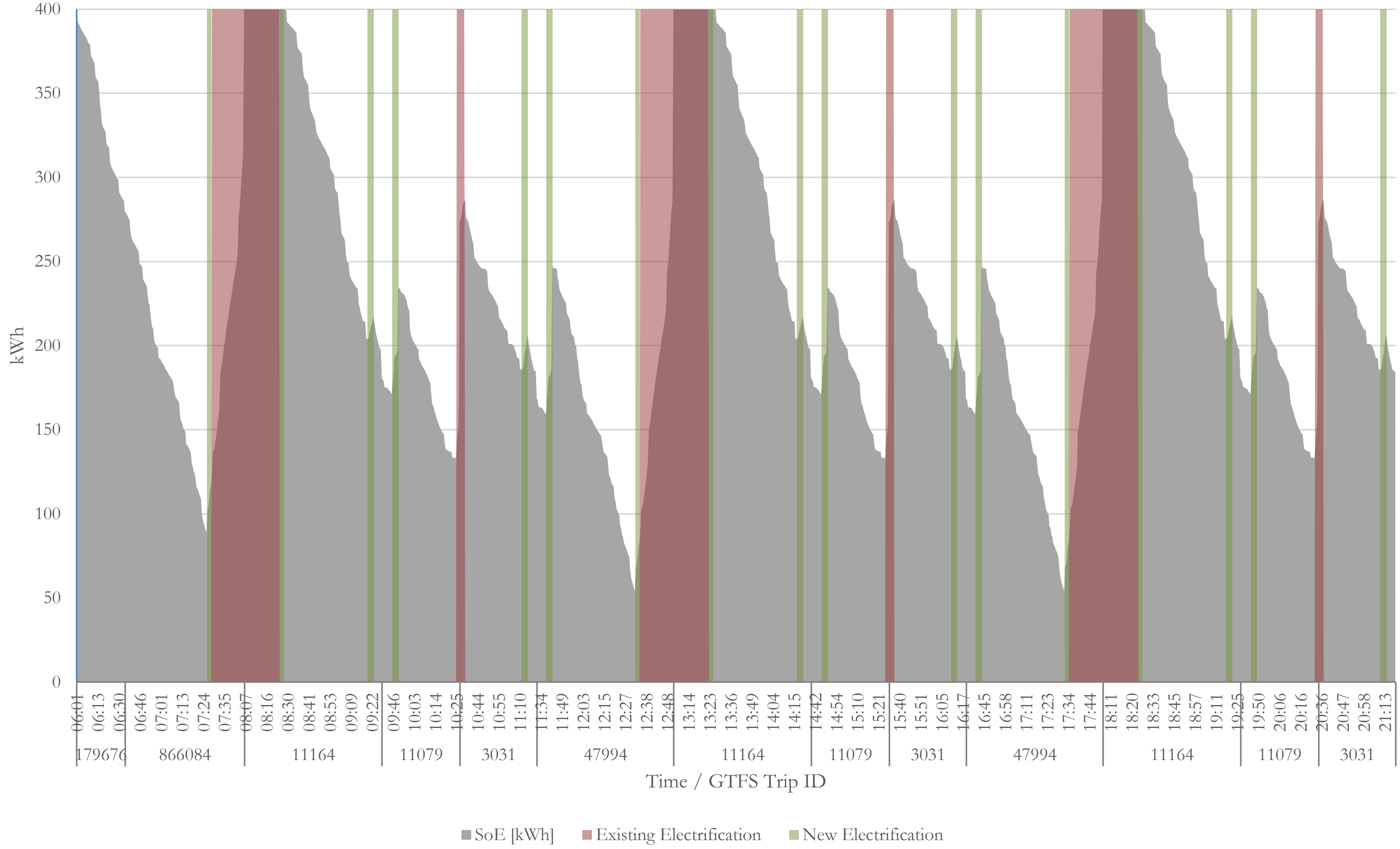

Figure 110: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run "expensive substation 1".

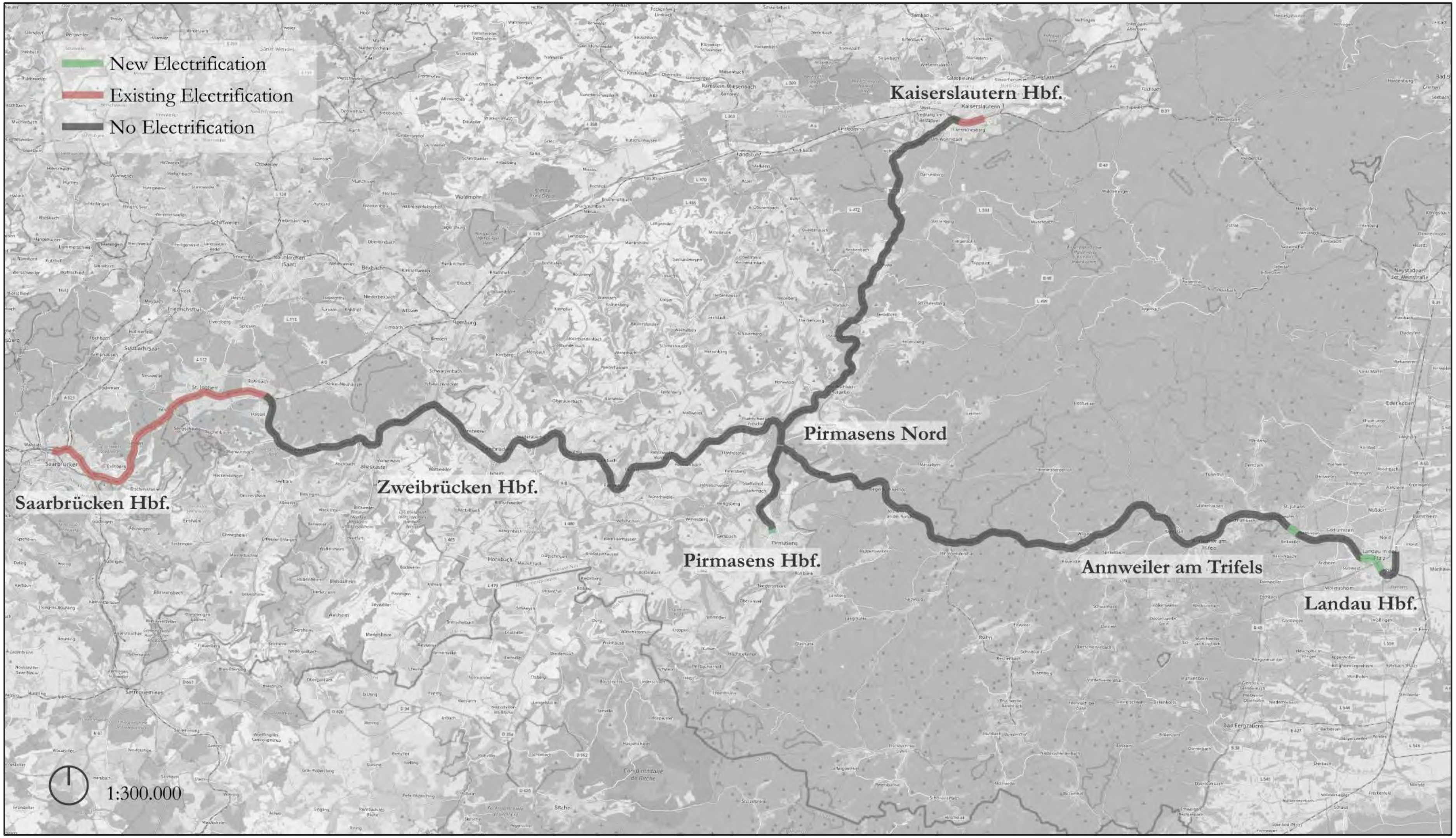

Figure 111: Existing and proposed electrification for model run “expensive substation 2”. The electrification is given per track even though this is not visible at map scale.

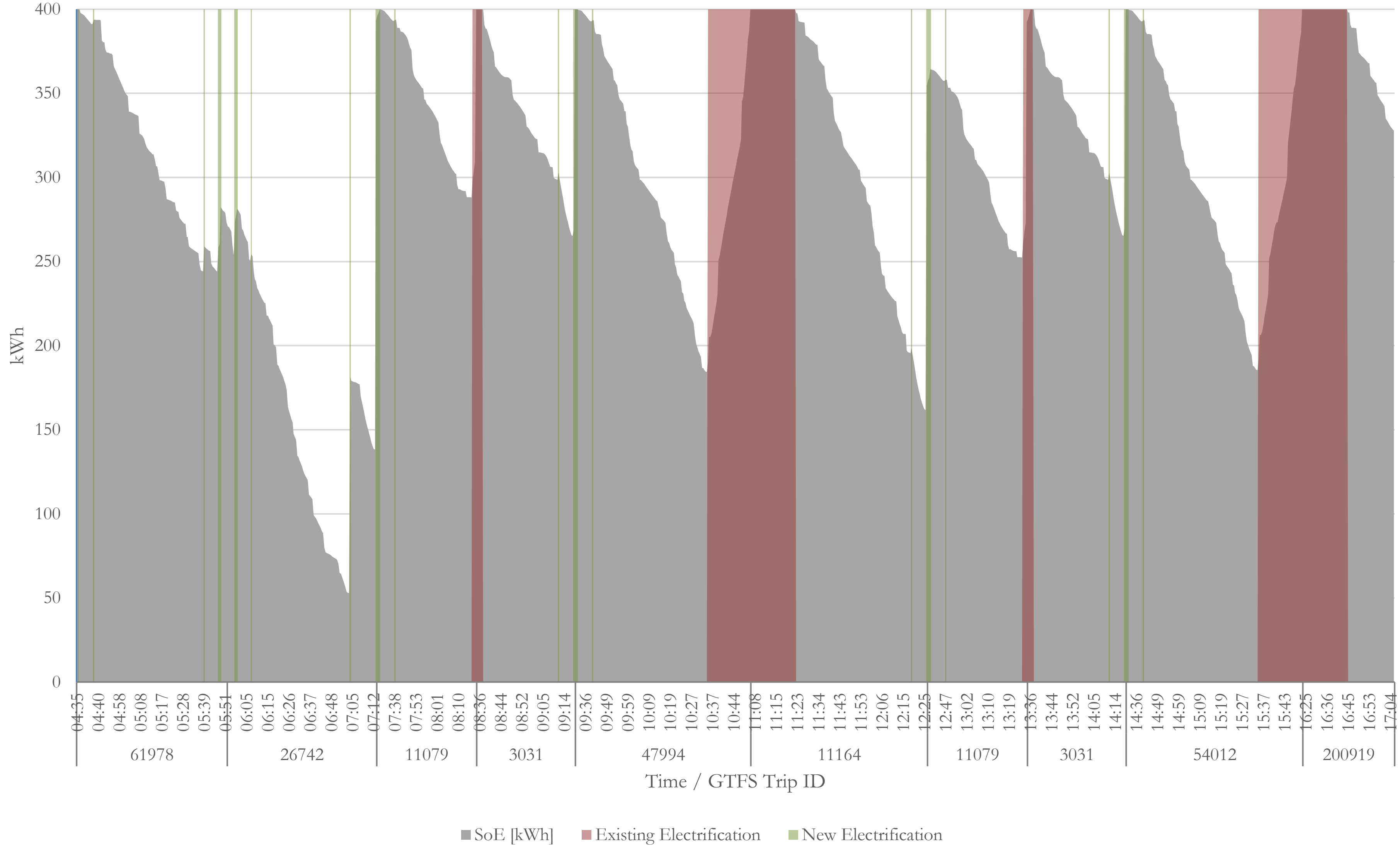

Figure 112: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run “expensive substation 2”.

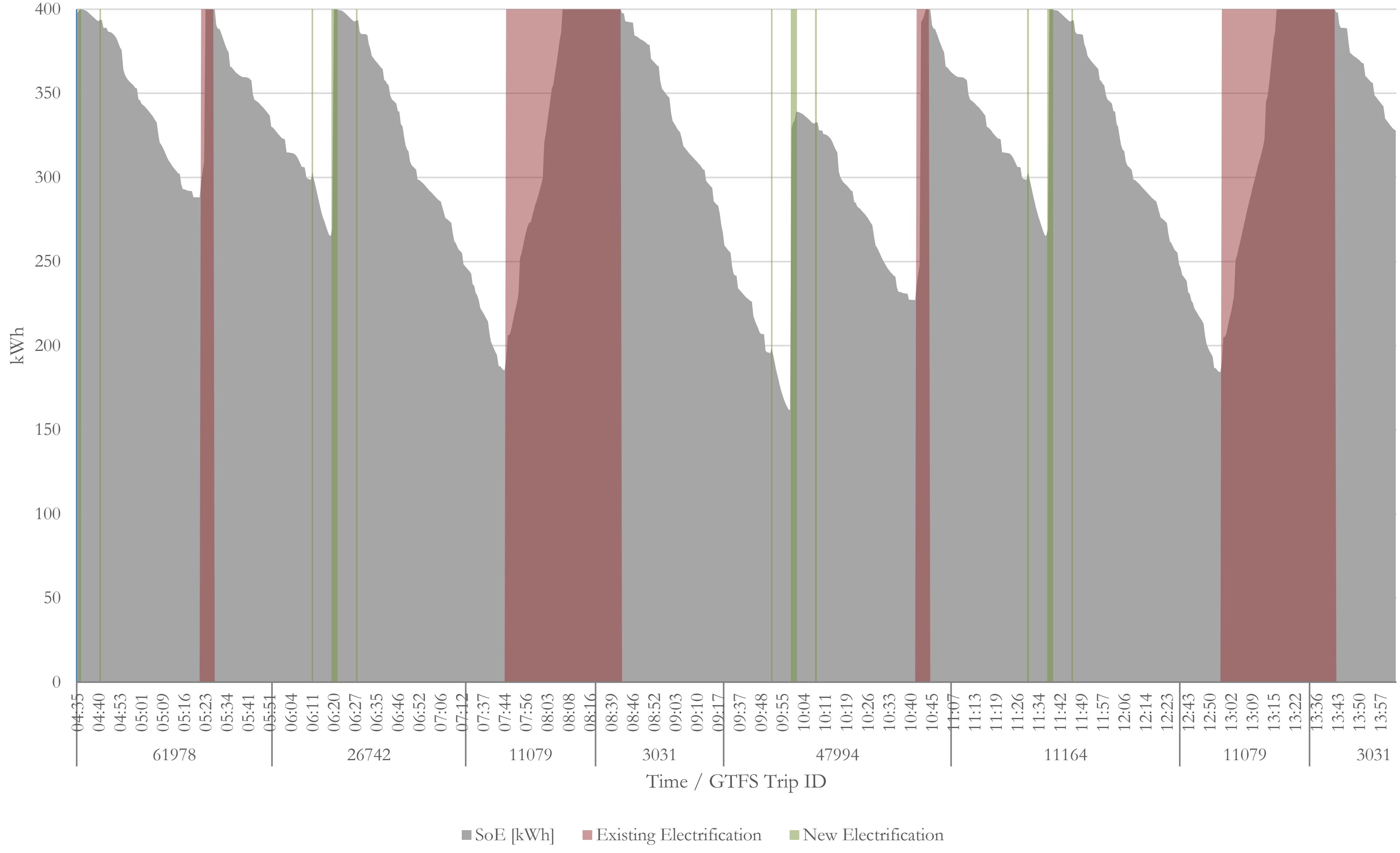

Figure 113: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run “expensive substation 2”.

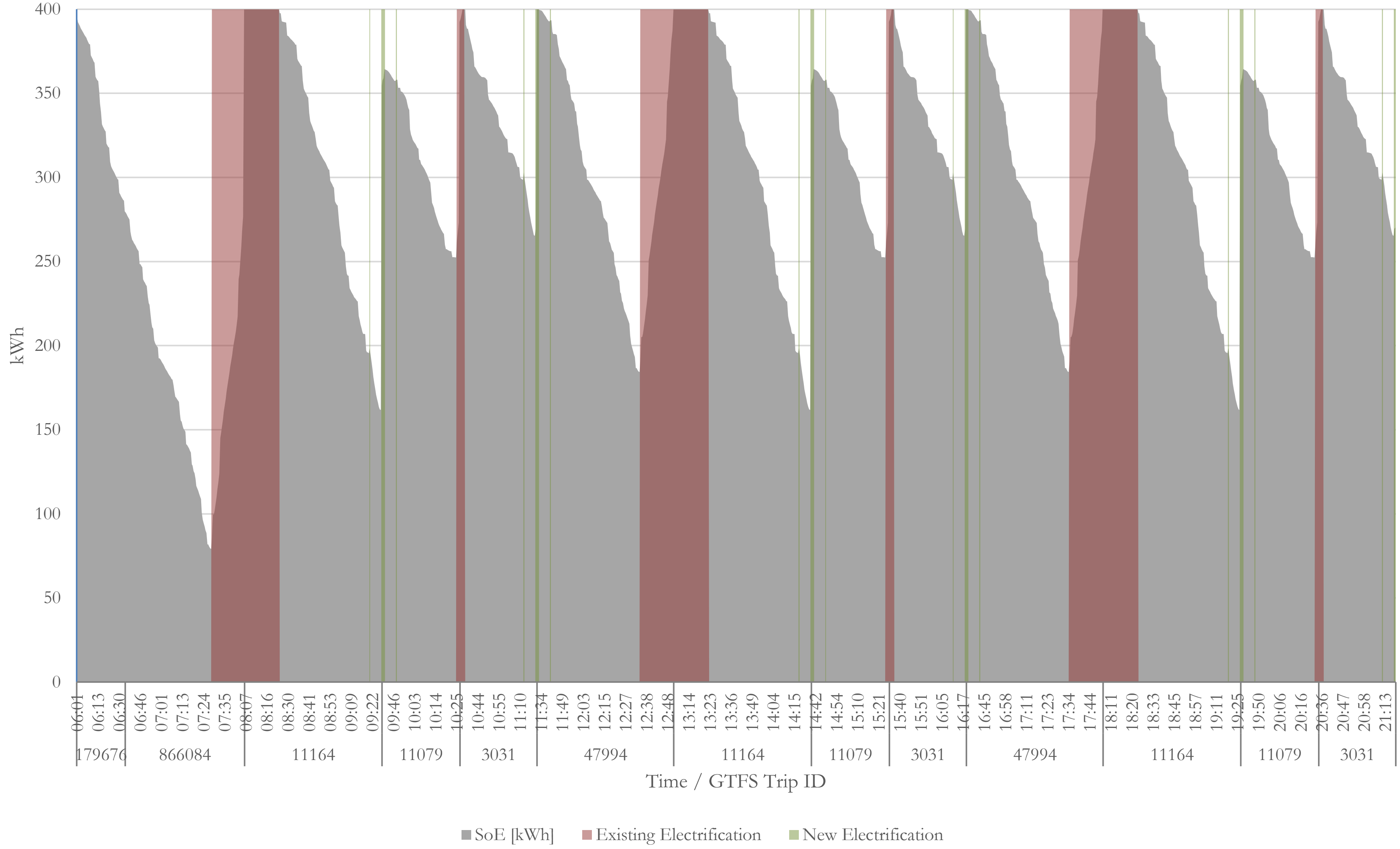

Figure 114: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run "expensive substation 2".

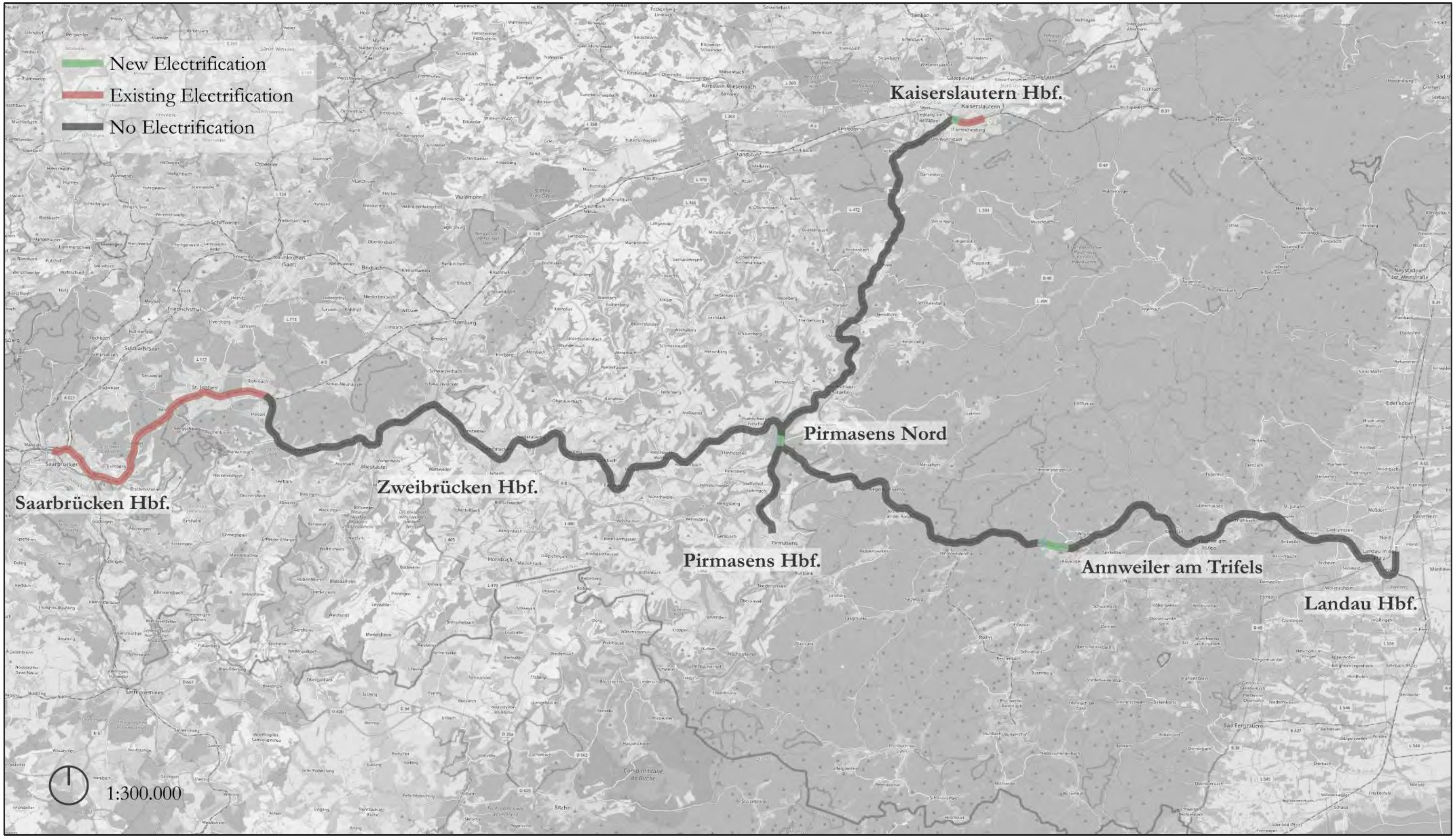

Figure 115: Existing and proposed electrification for model run "expensive substation 3". The electrification is given per track even though this is not visible at map scale.

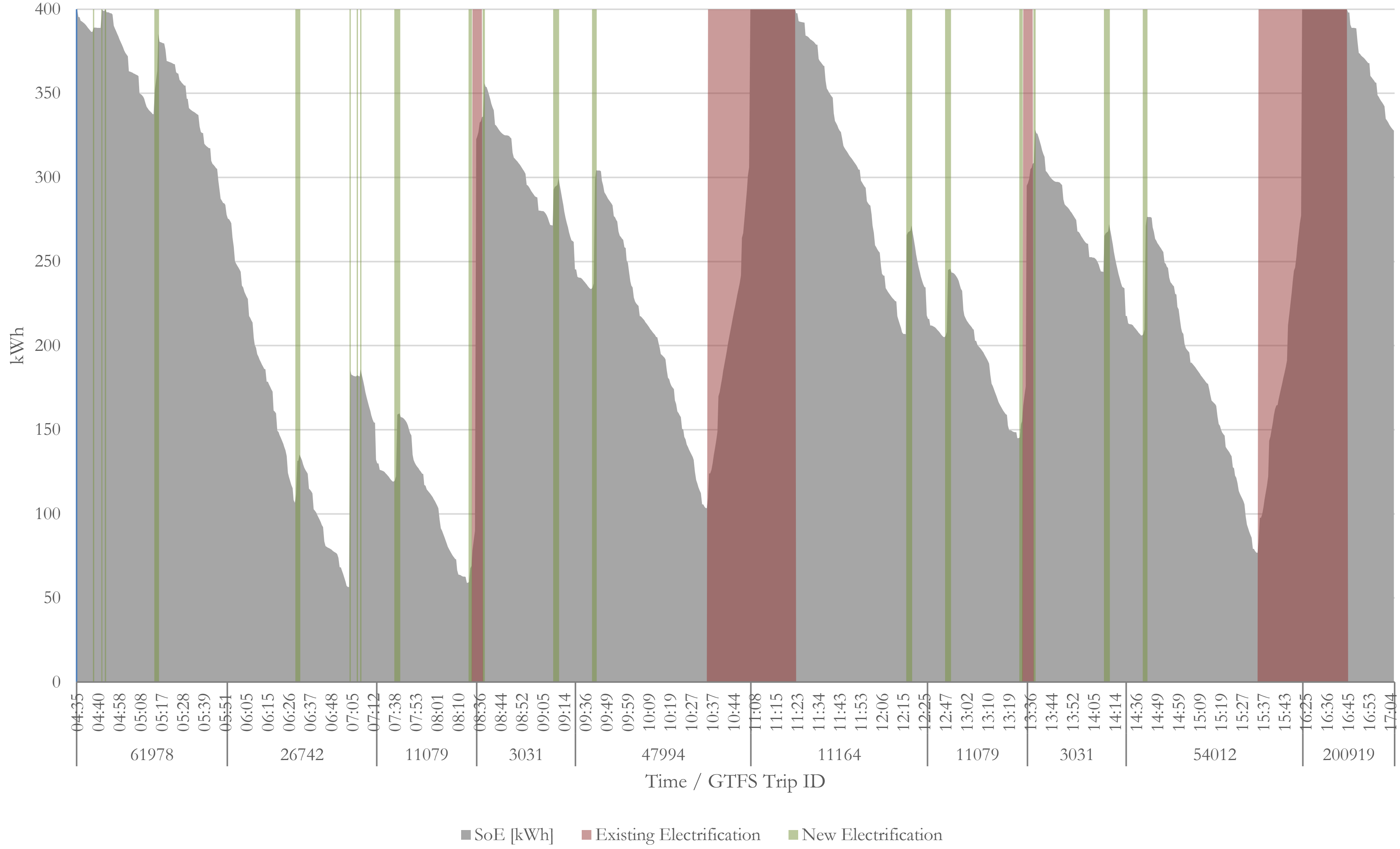

Figure 116: State of energy and the presence of electrification for vehicle 156632707 through its circulation given the distribution of existing and new electrification within the results of model run "expensive substation 3".

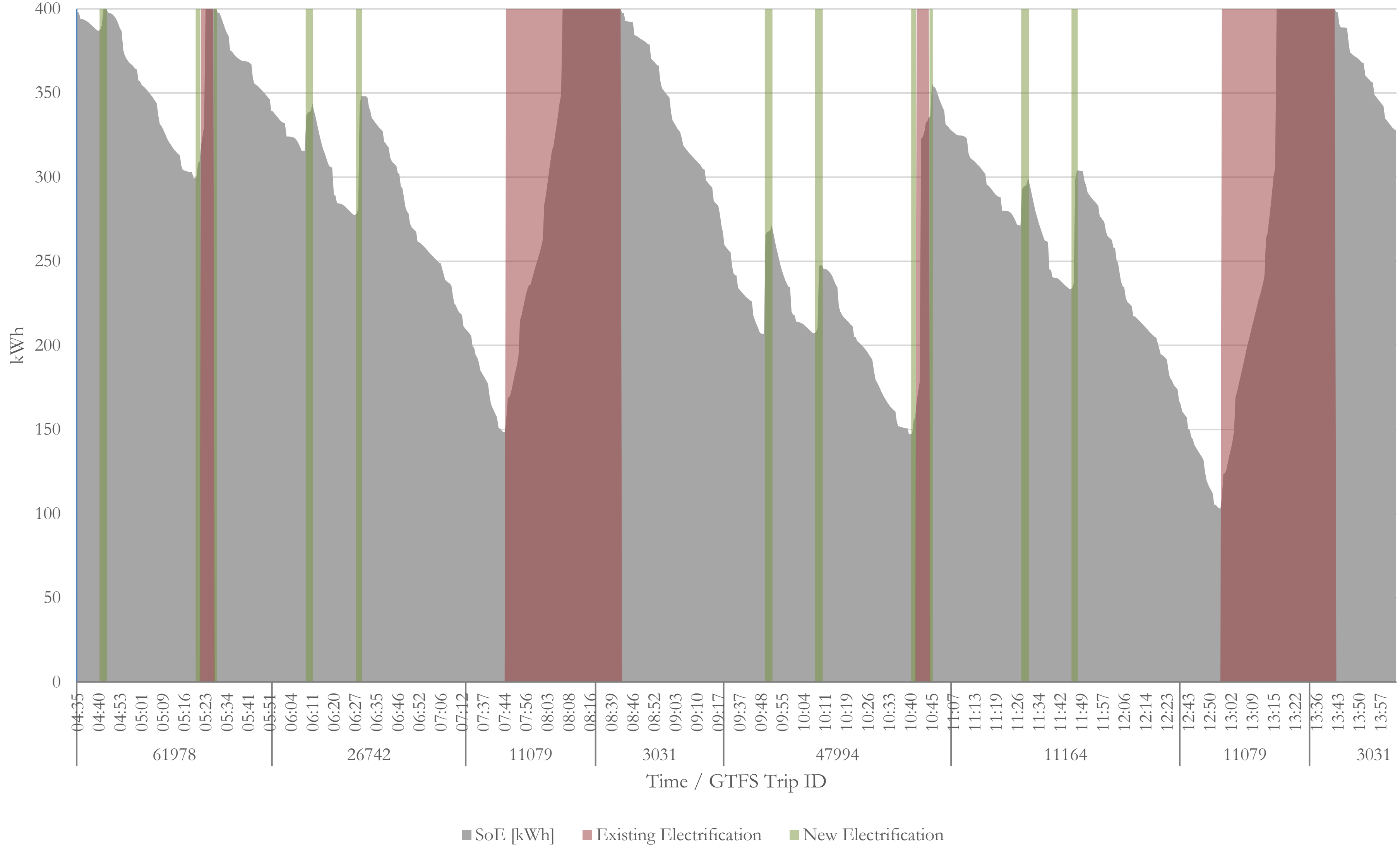

Figure 117: State of energy and the presence of electrification for vehicle 156632710 through its circulation given the distribution of existing and new electrification within the results of model run “expensive substation 3”.

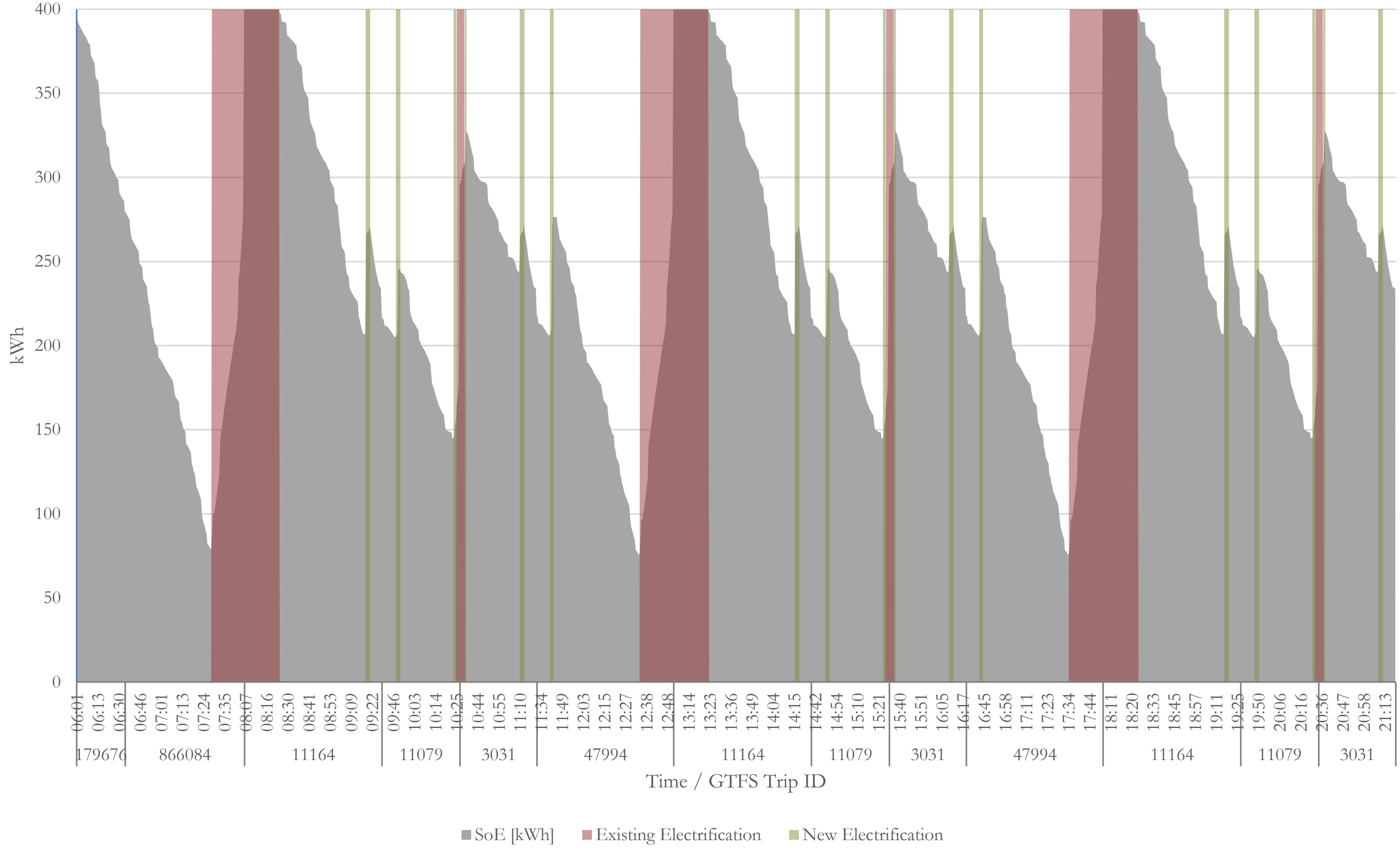

Figure 118: State of energy and the presence of electrification for vehicle 156632711 through its circulation given the distribution of existing and new electrification within the results of model run “expensive substation 3”.

11. Appendix C: General Map Visualizations

This appendix contains various map visualizations relevant to the understanding of the thesis contents. All maps are referred to and discussed within the thesis when relevant.

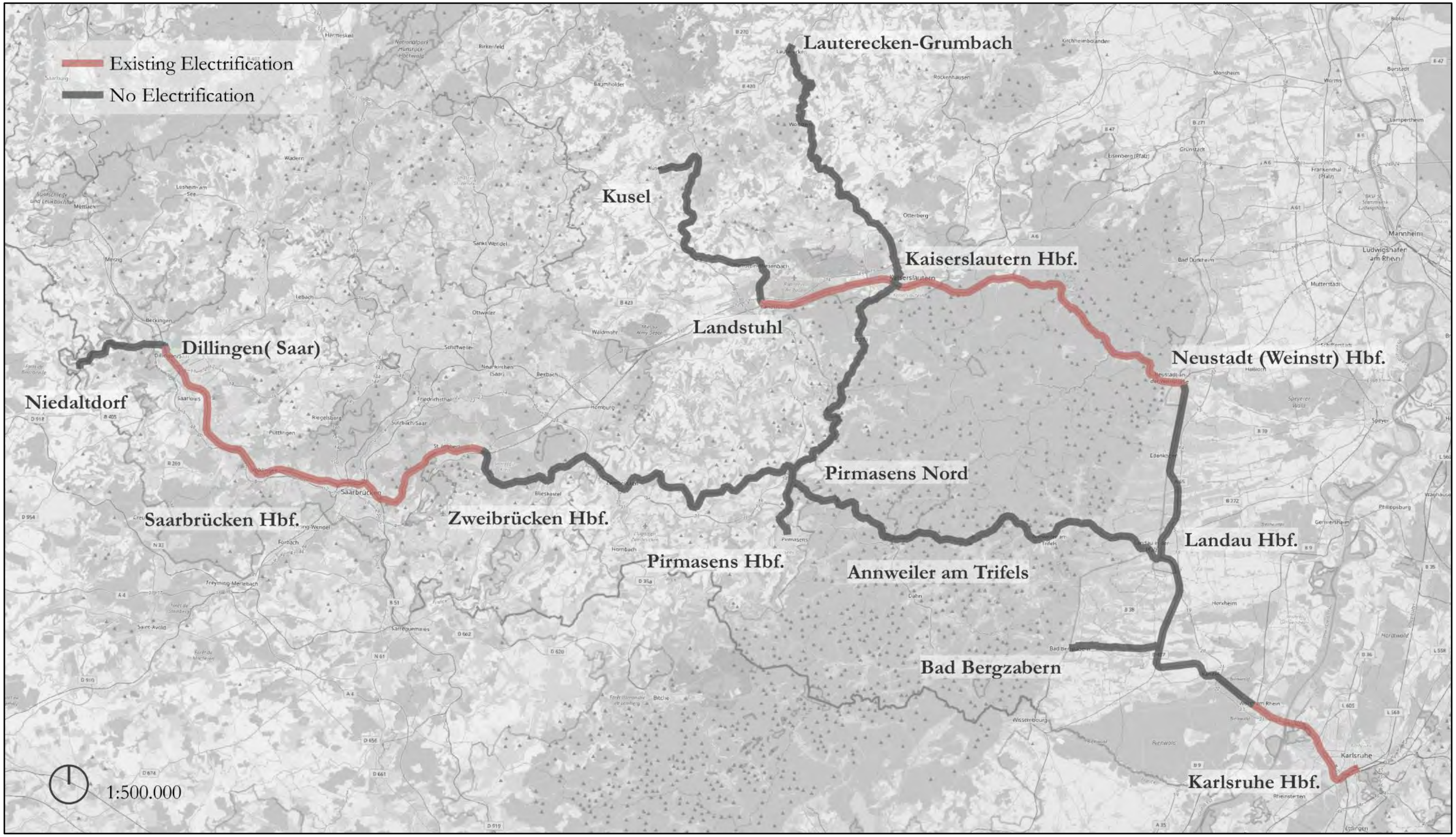

Figure 119: Electrification status of the entire network relating to the circulation of the vehicles in the edited Pfalznetz circulation plan.

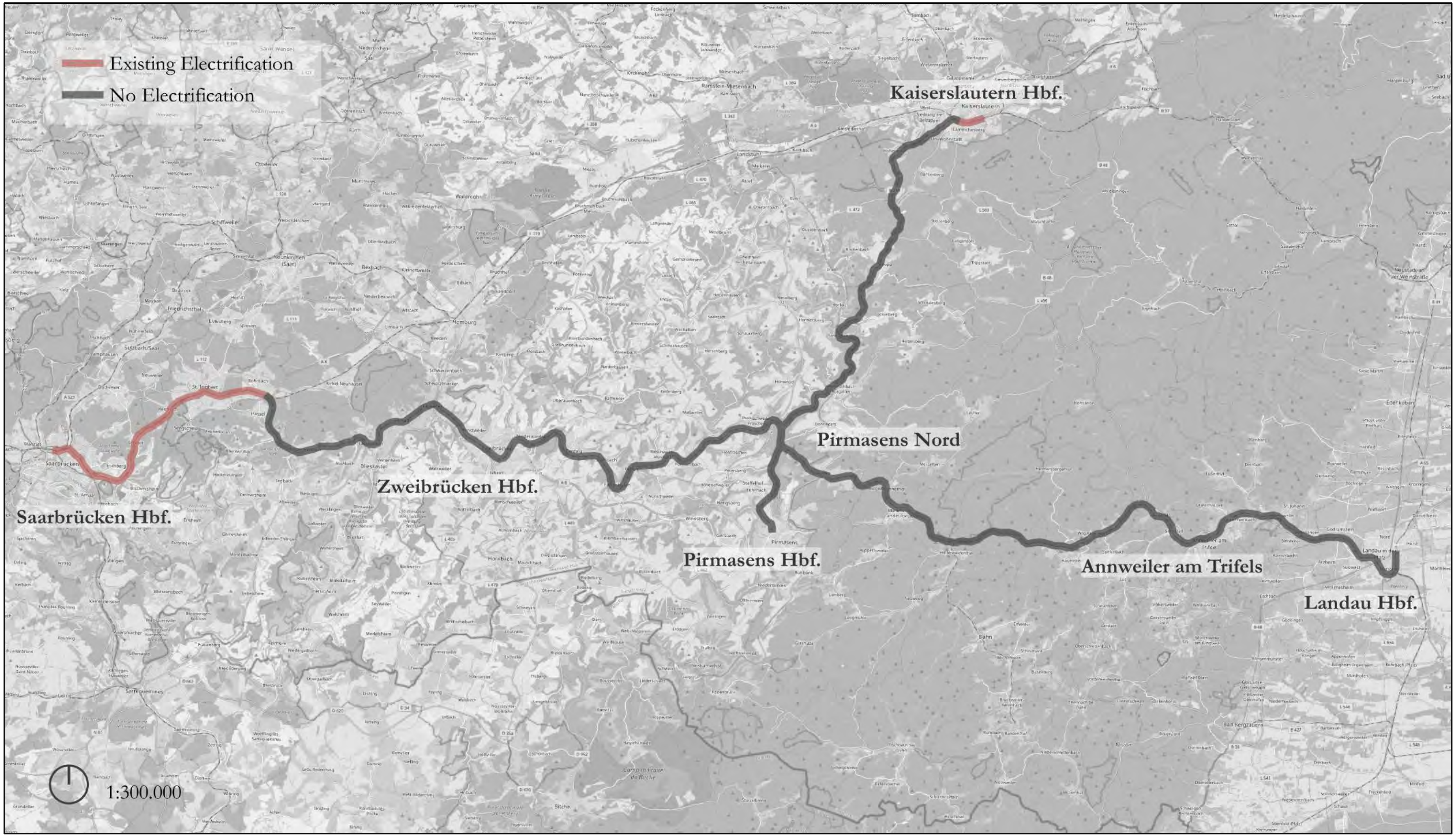

Figure 120: Electrification status of the network pertaining to the circulation of vehicles 156632707, 156632710 and 156632711.

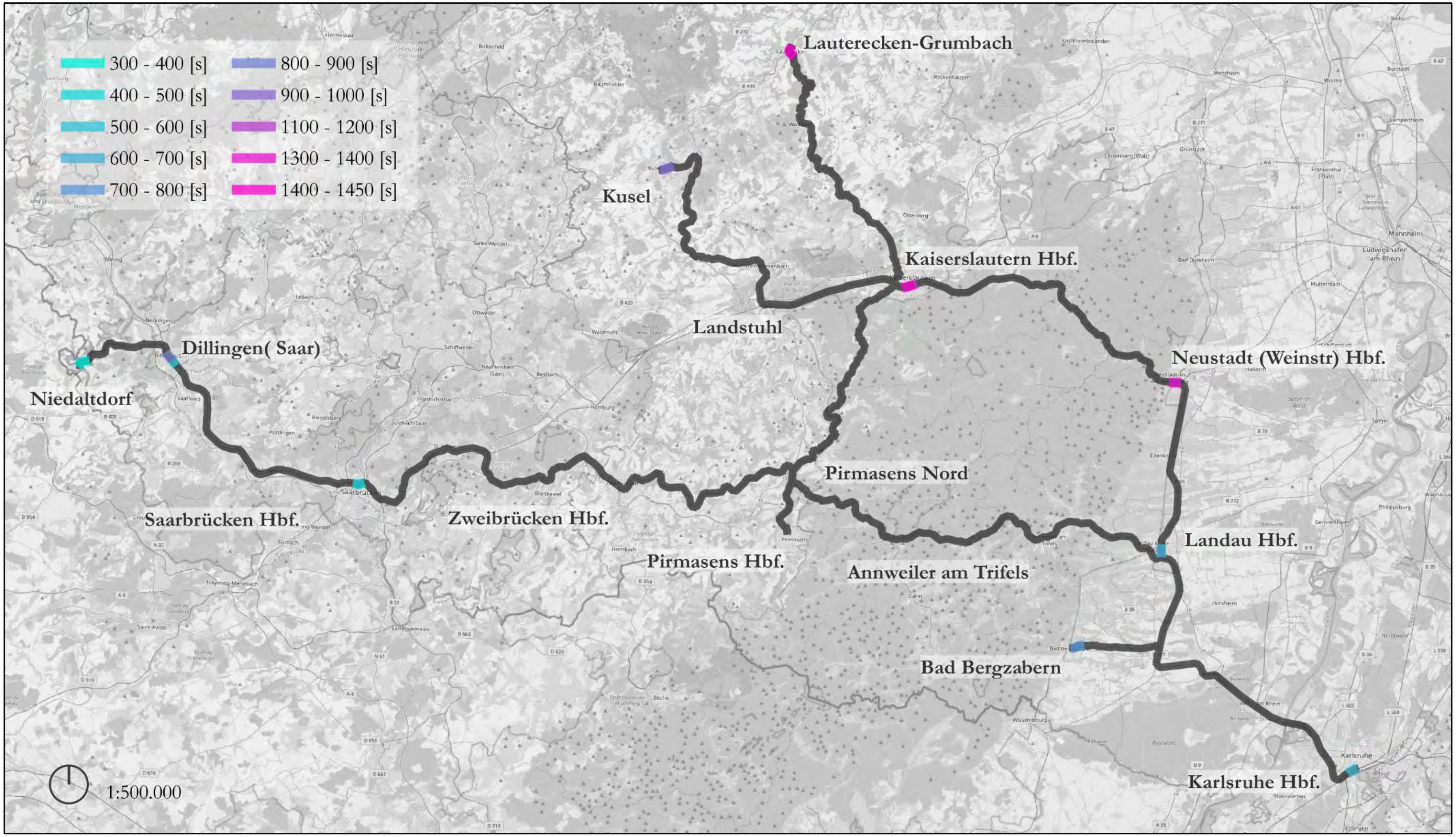

Figure 121: Map visualization of the average time spent at segments for each vehicle spending time at that segment. Only segments where the average time spent is above 5 min (300 seconds) are highlighted.

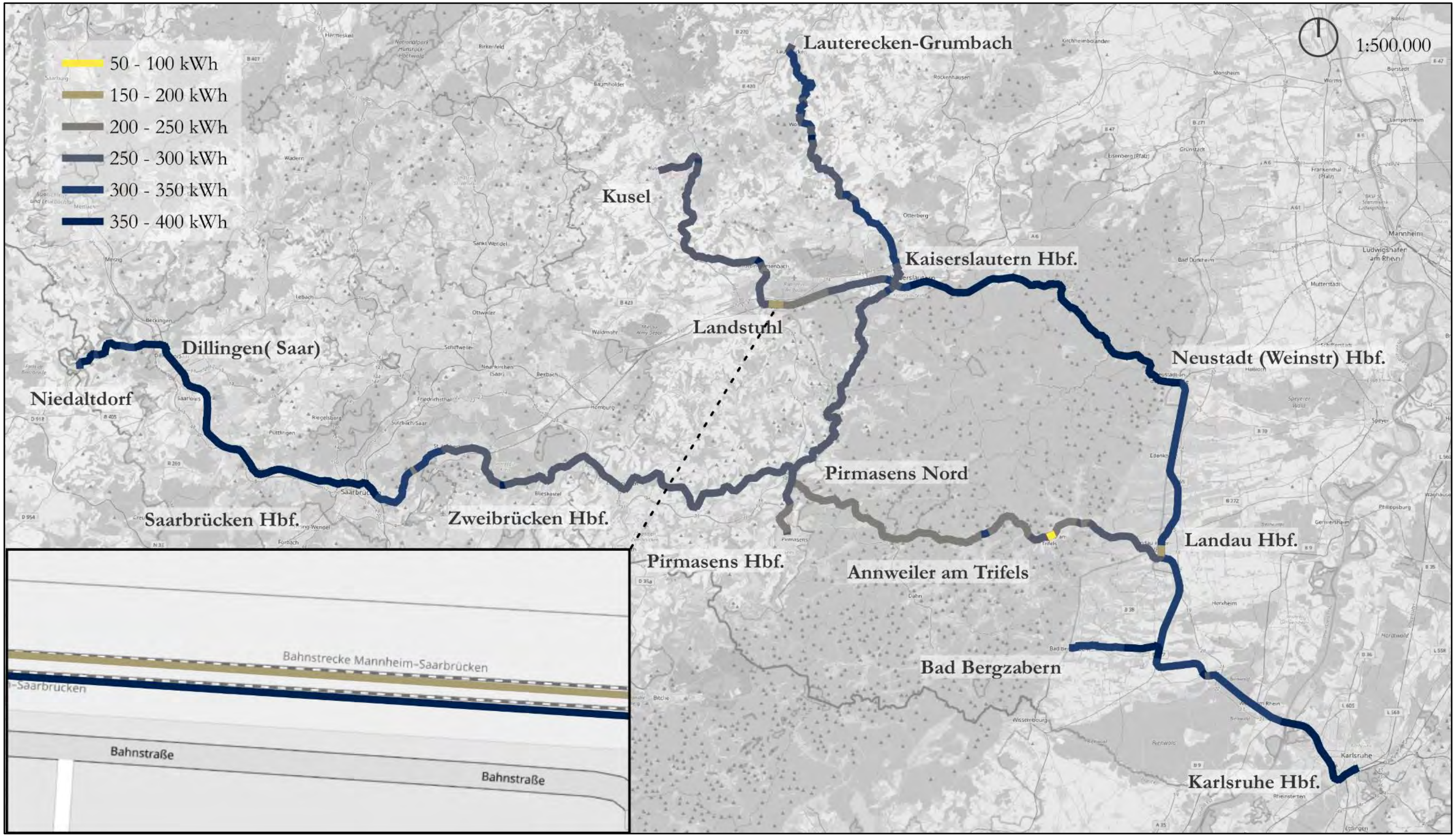

Figure 122: Map visualizing the average battery state of energy for each segment for all vehicles passing through that segment when they pass through it. A zoomed in section around Landstuhl is visualized to illustrate the by-track basis on which the optimization is conducted.

Jeder Abschlussarbeit ist auf der letzten Seite der Arbeit gesondert eine unterschriebene "Erklärung" beizufügen. Sie dokumentiert, dass die vorgelegte Arbeit selbst angefertigt und die aus fremden Quellen direkt oder indirekt übernommenen Inhalte als solche kenntlich gemacht wurden:

ERKLÄRUNG

Ich erkläre, dass ich die vorliegende Arbeit oder Teile davon nicht für andere Prüfungs- und Studienleistungen eingereicht, selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe. Sämtliche fremde Quellen inklusive Internetquellen, Grafiken, Tabellen und Bilder, die ich unverändert oder abgewandelt wiedergegeben habe, habe ich als solche kenntlich gemacht. Mir ist bekannt, dass Verstöße gegen diese Grundsätze als Täuschungsversuch bzw. Täuschung geahndet werden.

Berlin, den 28.09.2023

Unterschrift

[redacted]