agriculture

MDPI

Article

Road Freight Demand Forecasting Using National Accounts' Data—The Case of Cereals

Taha Karasu [1,*], Pekka Leviäkangas [1] and David John Edwards [2,3]

[1] Department of Civil Engineering, University of Oulu, 90570 Oulu, Finland; pekka.leviakangas@oulu.fi
[2] School of Engineering and Built Environment, Birmingham City University, Birmingham B4 7XG, UK; drdavidedwards@aol.com
[3] Faculty of Engineering and the Built Environment, University of Johannesburg, Johannesburg 2092, South Africa
* Correspondence: taha.karasu@oulu.fi; Tel.: +358-50-406-7589

Abstract: This paper investigates the potential of utilising historical agricultural production data for enhancing road freight transport forecasting, focusing on cereal production. This study applies a multiple linear regression analysis using national statistical accounts and secondary data. The data were sourced from Finland's Statistics Agency and the Natural Resources Institute. The analysis identifies an observable correlation between agricultural production and road freight volumes, although this correlation is not statistically significant. The highest adjusted R^2 observed in the models was 0.62. The analysis reveals that previous years' production data can help forecast future road freight volumes, with vehicle mileage estimable from recent production and stock levels. Additionally, annual percentage changes in the volume of transported cereals can be partially predicted by the changes in total available cereals and opening stocks from two years prior. This exploratory research highlights the untapped predictive potential of agricultural production variables in forecasting road freight demand, suggesting areas for further forecasting enhancement.

Keywords: supply chains; demand forecasting; road freight; agriculture; regression; cereals

check for updates

Citation: Karasu, T.; Leviäkangas, P.; Edwards, D.J. Road Freight Demand Forecasting Using National Accounts' Data—The Case of Cereals. *Agriculture* **2024**, *14*, 1980. https://doi.org/10.3390/agriculture14111980

Academic Editors: Karel Tomšík, Pavel Kotyza and Josef Abrhám

Received: 4 October 2024
Revised: 30 October 2024
Accepted: 1 November 2024
Published: 5 November 2024

1. Introduction

Processing and transportation together consume approximately one-third of the total energy used in agricultural production [1]. Despite this energy expenditure, inefficiencies in the agricultural supply chain led to the loss of more than a third of all produce globally [2,3]. Moreover, the trends of globalisation and the increasing size of farms, coupled with a decrease in the number of farms, have extended the average distance between producers and consumers [4,5]. Cumulatively, these socio-economic changes underscore the importance of accurately estimating transport demand to contribute to productive and sustainable agricultural transportation [6,7]. Additionally, the efficiency of agricultural transportation varies greatly by location, indicating the need for localised efforts to enhance transportation efficiency [8,9].

More accurate and timely forecasting of freight demand is integral to transport planning and management, contributing to the advancement of sustainable and efficient logistics [10]. Many European countries, such as Sweden, Norway, and Finland, have developed freight forecasting models like Samgods, the Norwegian Freight Transport Model, and the TRIMODE model used by the European Commission [11–15]. These models predominantly focus on overall transport patterns and lack specific dynamics of the agricultural sector, for example, volumes of production and stocks. To date, the relationship between transportation parameters and production parameters from previous years remains unexplored. Typically, continuous trends are extrapolated while taking partially into account how the industrial architecture is changing—for example, if there is a decline in the production of a particular type of commodity.

Although studies have explored transport demand forecasting in other sectors [10,16–19], research focusing on freight forecasting in the agricultural domain is limited. For instance, a study by Miklius et al. [6] examined transport demand forecasting for agricultural commodities, yet does not fully account for recent advancements in logistics or the modern agricultural landscape. To the best of our knowledge, there is no recent study that integrates agricultural production data with road freight forecasting models. This creates a discernible research gap, particularly in the domain of forecasting agricultural transport demand based on production volumes.

This study contributes to filling this prevailing knowledge gap by exploring the case of cereals as a product group. Road freight transport is focused upon because it is the most utilised mode in the agricultural sector [20]. Insights will be provided into what could (or could not) be forecasted based on presently available data. Specifically, the research explores the possibility of introducing an additional dimension to existing freight forecasting models utilising production data from national accounts. Through a comprehensive understanding of the quantity and characteristics of transportation variables associated with production data, the study contributes to optimising logistics networks within the transportation industry. It also contributes to determining appropriate infrastructure investments and efficiently allocating resources. By employing a quantitative research design (specifically, multiple linear regression), this study uncovers patterns and correlations between agri-production and road freight volumes. Through the analysis of publicly available data from the Statistics Agency of Finland and the Natural Resources Institute of Finland, the methodology adopted ensures transparency and replicability. Ultimately, work will enhance the predictability of road freight demand and understanding of the dynamics between agricultural production and transport need.

2. The Case of Cereals

In Finland, cereals provide a cornerstone for domestic food, utilised directly by humans, as domesticated animal feed and industrial input, notably in beverage production [21]. Finland's cereal harvest has fluctuated between 2.6 and 4.2 million tons since 2000 [22]. Along with the amount of harvest, inventory levels, market conditions, and the proportion of use of cereals through different channels might be associated with transportation-related indicators of future years. Within the category of 'cereals', encompassing barley, oat, wheat, and rye, barley stands out as the most extensively cultivated grain [21,22]. Historically, a substantial portion of barley, accounting for 65% of domestic consumption, has been allocated to livestock farming. However, this proportion is showing a decreasing trend in recent years. Cereals produced domestically are predominantly consumed within Finland [22]. However, in certain years, heightened import levels are noted, especially during adverse harvesting seasons [22]. Notably, cereals in Finland are often stored on farms rather than in industrial warehouses [21]—an idiosyncrasy that reflects the local approach to agricultural storage within the Finnish context.

Figure 1 illustrates the cereal supply chain with a focus on transportation pathways. Excess harvest is typically stored at nearby facilities rather than being immediately transported to industrial warehouses [21]. Cereals from agri-farms are used on site, sent to livestock farms as feed, and transported to industrial facilities for processing to be used for food [22]. The supply chain extends to merchandise warehouses that handle the cereals before distribution. International transport hubs also feature in the diagram, signifying the cereals' journey into global markets. In the diagram, each connection point with a 'truck' symbol highlights a juncture for applying transport demand forecasting.

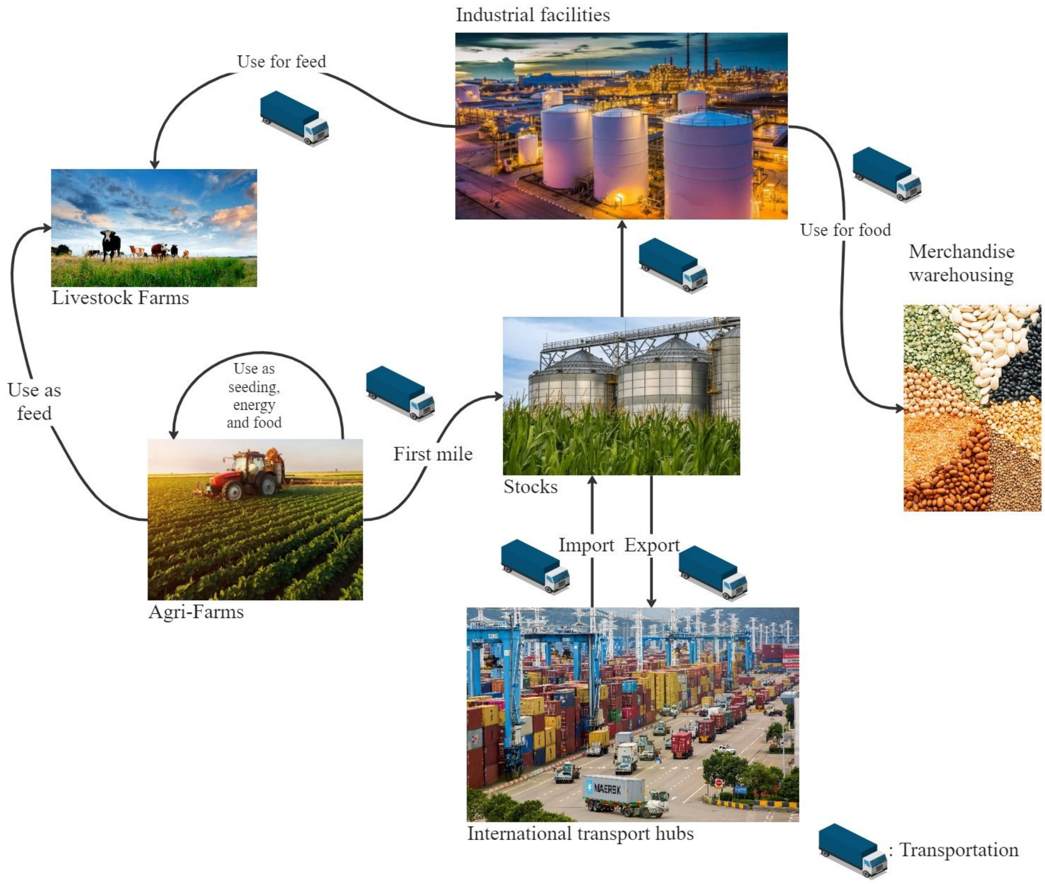

Figure 1. Road freight demand within the national accounts' system showing the flows of the commodity (each block represents a dedicated accounting line in the system).

3. Freight Transport Forecasting

The literature highlights the role of freight transport forecasting across multiple domains [16,18,19,23]. There is a particular focus on improving transport efficiency, reducing environmental impacts, and supporting sustainable growth in agriculture and other industries [16,17]. Some forecasting models use a time series analysis and/or econometric modelling to predict demand based on historical data [17,18,23]. Jayanthi and Jothilakshmi provide a comprehensive survey on time series forecasting, particularly in intelligent transport systems, categorising various methods and their outcomes to aid practical applications in traffic management [17]. Meanwhile, studies like Petrik et al. [16] and Bilbao-Ubillos et al. [18] underscore the importance of adapting forecasting methodologies to capture demand patterns and non-linear dynamics. Norouzian-Maleki et al. [19] combine system dynamics with a data envelopment analysis, offering a dual framework that forecasts demand and guides optimal policy decisions. Kii et al. [23] contribute by focusing on land use and transport interaction models, advocating for hybrid approaches that address sustainable goals. Together, these studies point out the growing recognition of context-sensitive forecasting models, where approaches are tailored to distinct transport sectors. However, despite these studies, there is limited literature addressing transport demand forecasting specifically within the agriculture sector.

Forecasting transport demand is quintessentially important to freight transport planning. It influences the formulation of development policies and infrastructure investment and serves as a key metric for assessing planning rationality [16]. Capturing robust correlations in both spatial [23] and temporal [17] aspects among various transportation demands can substantially enhance the accuracy of demand forecasting [18,19].

National Models for Freight Transport Forecasting

In Finland's framework for managing freight logistics and environmental impacts within the maritime and transit sectors, specific models such as FRISBEE, MERIMA, and TRAMA have been developed and are employed in Finland [11]. These models are designed for distinct purposes but with complementary functions. They aim to support the decision-making process of regulators, thus contributing to the strategic development of Finland's transport infrastructure and environmental policies. The FRISBEE Model is employed to allocate import and export transport across various transport routes and modes within predefined scenarios [24]. FRISBEE facilitates the analysis and optimisation of transportation logistics, enabling informed decision making regarding the allocation of goods to different routes and modes based on specified criteria. MERIMA is designed to compute emissions originating from international maritime transport [25]. It offers an in-depth analysis across different scenarios, encompassing various levels of assessment including total maritime transport, port pairings, routes, and vessel calls. This enables the scrutiny and evaluation of emissions, facilitating informed decision-making processes in maritime environmental management. TRAMA serves as an analytical instrument dedicated to computing both the economic ramifications and emissions associated with transit transport [11]. It accommodates a broad spectrum of scenarios, thus enabling assessments of transit-related economic impacts and emissions across various contexts [26]. TRAMA consists of two computer models for estimating, monitoring, and analysing economic effects of transit freight transport. As delineated in the report from the Ministry of Transport and Communications, Finland [11,26], the model is intended for use by authorities. It aids in the evaluation of the economic benefits and detriments of transit traffic to Finland, as well as its employment effects. Notably, the forecast for goods transport, based on the transport intensity method tied to Finland's production structure, projects that lorry transport performance will increase by 8.3% from 2021 to 2030. Continued growth is expected until 2040, followed by a slight decline [27]. This underscores a future shift in road freight volumes, which although not directly linked to agricultural variables, impacts overall transport resource allocation and environmental strategies. While existing freight forecasting models (viz. FRISBEE, MERIMA, and TRAMA) offer insights into the allocation of transport resources and environmental impacts in Finland, they primarily focus on overall transport patterns rather than specific correlations with agricultural production variables [11,24,26].

Neighbouring countries such as Sweden, Norway, and Denmark and initiatives at the European level utilise distinct freight transport models that are tailored to their specific contexts and needs. At the European level, the TRIMODE model (developed by the European Commission) functions as a comprehensive simulation tool. It is designed to analyse and assess Europe's intricate interplay between transport, economic, and energy systems, with a specific focus on evaluating the impacts of large-scale transport infrastructure projects [12]. In Sweden, the national Samgods goods flow model consists of two separate sub-models, i.e., the logistics model; and the railway capacity model [13]. The logistics model divides various types of goods into transport chains, thereby minimising logistics costs, while the railway capacity model assesses railway capacity constraints [13]. In Norway, the freight transport model comprises demand and supply components. The demand aspect encompasses goods flow matrices between Norwegian municipalities and between Norway and other countries [14]. Similarly, Denmark utilises a comprehensive model to analyse total traffic flows within its borders and internationally. This model can calculate incoming traffic based on user-selected assumptions [15].

4. Methodology

This study is primarily grounded in positivism and pragmatism, emphasising the use of objective, measurable data to explore the prospective patterns and relationships to generate workable predictions of the phenomenon under investigation [28]. The application of a statistical analysis and the deductive approach in defining variables and exploring potential relationships align with the positivist tradition [29–31].

4.1. Methods and Research Process

The research approach is designed accordingly as illustrated in Figure 2. The research commenced with the precise formulation of research questions.

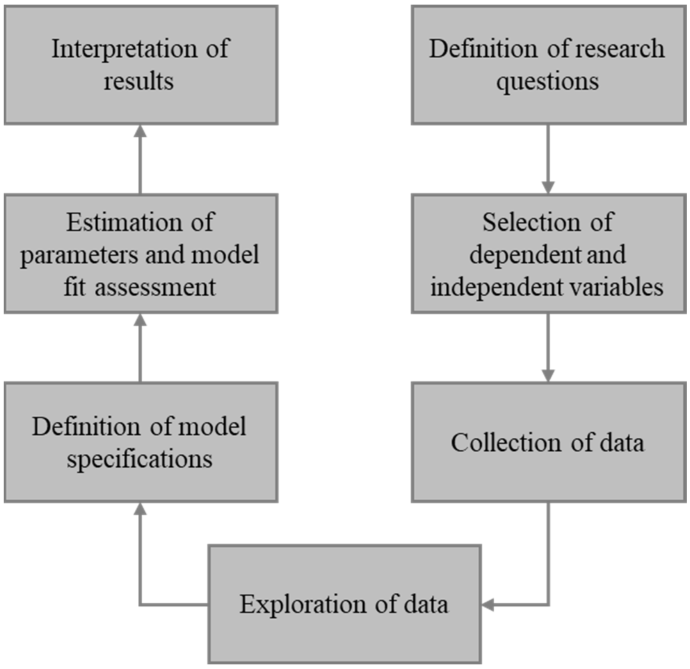

Figure 2. Research process.

RQ1: What constitutes the available data, and which variables are pertinent for road freight forecasting?

RQ2: What is the logical association between input variables and forecasted variables?

RQ3: To what degree can the road freight transport volumes be forecasted?

Corresponding to these questions, the selection of dependent and independent variables was undertaken, followed by the compilation of annual data. Initially, public databases from the Statistics Agency of Finland (for transportation-related data) and the Natural Resources Institute of Finland (for agricultural data) were scanned. To ensure an understanding and comprehensive collection of data, direct communication with both institutions was established. Specifically, the study focuses on dependent variables directly related to the transport of cereals. These include the vehicle mileage of cereal-carrying trucks, volume of goods representing the amount of cereals carried by these trucks, and volume of transport, which is determined by multiplying the vehicle mileage when the trucks are loaded with cereals and the actual amount of cereals in the trucks. Independent variables encompass agricultural production metrics.

Despite the dataset's temporal scope being seemingly concise, it is important to note that this study leverages the most comprehensive and consistent data available from the periods 2009/2010 to 2021/2022 for agricultural production data, and from 2011 to 2022 for transportation data. The selection of this specific timeframe was dictated by the availability of reliable and methodologically consistent data. Given the historical methodological changes in data collection and the subsequent inconsistencies in the data recorded before these periods, the dataset employed herein represents a more viable option for conducting this research. Within the constraints of data availability, this dataset is considered sufficient for the study's objectives. It provides a valuable, albeit limited, perspective on the dynamics at play within the specified periods. This allows for an exploration of trends and relationships with the acknowledgment that the scope is determined by the data's accessibility and reliability. Thus, while acknowledging the limitations posed by the available data, this study proceeds on the premise that the utilised dataset is the most suitable and comprehensive for addressing the research questions posed, within the constraints of existing data collection practises. This approach underscores a pragmatic adaptation to the available resources, ensuring that the research is grounded in the best possible empirical evidence under the circumstances.

In addressing *RQ1*, the authors engaged research specialists at the Natural Resources Institute of Finland [32] and Finland's National Statistics Institute [33] to collate existing data and conduct interviews with researchers from these institutions to clarify data-related queries. For *RQ2*, the study constructed a diagram illustrating the interrelationship between data elements through a transportation-focused perspective. In tackling *RQ3*, the investigation delves into the efficacy of potential models while scrutinising the outcomes derived from these models.

Regression models were chosen for their robustness in identifying and interpreting the relationships between multiple variables, aligning with the research goal to explore the impact of agricultural production metrics on agricultural transportation [34]. This approach facilitates an analysis of how changes in independent variables (e.g., cereal production) influence dependent transportation variables, including the vehicle mileage, volume of goods, and volume of transport. Unlike autoregressive integrated moving average (ARIMA) models, which are better suited for univariate time series forecasting within stationary datasets, a regression analysis can accommodate the interactions between multiple predictors and their impacts on transportation metrics [35]. This capability makes a regression analysis particularly relevant for this present research, offering insights into the multifaceted dynamics between the agriculture and transportation sectors.

A multiple regression model involves two or more independent variables and is mathematically expressed as

$$y_t = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \quad (1)$$

where y is the variable to be forecasted; and $x_1, \ldots x_k$ are the k number of independent variables. $\beta_1, \ldots, \beta_k$ measure the effect of each independent variable after taking into account the effects of all the other independent variables in the model. Thus, the coefficients measure the marginal effects of the independent variables.

Before constructing the models, prerequisites for their validation were established. Recognising the penalising effect of excessive variables, adjusted R^2 was favoured over R^2 [34], with an expected threshold set at above 0.55. To ensure validity, collinearity among independent variables was monitored using the variance inflation factor (VIF), and co-variants were excluded from the models. A significance limit of a p-value at 0.05 was established. In summary, a model is deemed valid if its adjusted R^2 surpasses 0.55, its p-value is <0.05, and the model exhibits no covariance, suggesting a potential causal relationship between independent and dependent variables. This validity should be routinely assessed with updated data. Given the regression analysis's objective to examine the potential influence of past agricultural data on transportation-related data, it is imperative that independent variables encompass prior inputs. Consequently, for a given transport year (T), agricultural inputs are derived from the agricultural season of T/T − 1 (specifically for production) and the agricultural season of T − 1/T − 2. Models were built using IBM SPSS Statistics 29.0 and R (Version 4.3.0) software.

4.2. Variables

The agricultural season commences in July of each year and extends through the end of the following June [22], and the harvesting season for cereals is limited to the months of August and September of each year [21]. However, transportation statistics are reported on a calendar-year basis [36]. This implies that statistics regarding the transportation of cereals are disclosed from January to December, while agricultural season statistics span from July to June (see Figure 3).

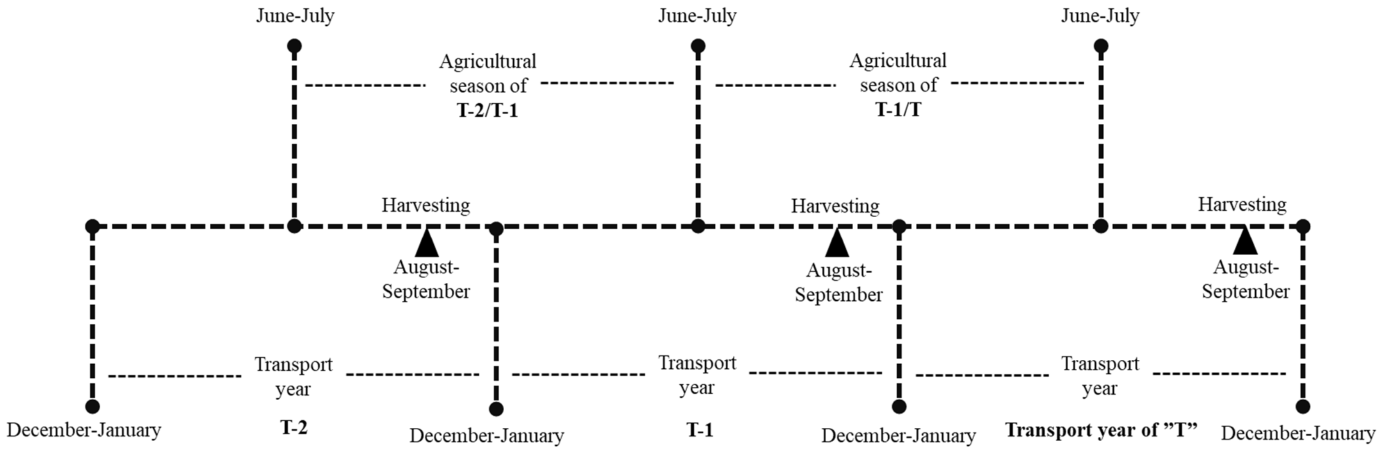

Figure 3. Agriculture production and transportation data on a timeline.

While the literature provides limited information on specific forecasting models in the context of agricultural transport, the importance of accurate transport demand forecasting has been widely acknowledged [37–40]. The data about the transportation of cereals were collected from Finland's national statistical institute (see Table 1) [33]. From 2011 to 2022, the data were extracted from one source [36]. For previous years, data availability varies across different locations. Methodological discrepancies in data collection prior to 2011 led to the exclusion of these earlier datasets from the analysis. The variables included as dependent variables for a further analysis are the vehicle mileage (measured in 1000 km) of cereal-carrying trucks, volume of goods (cereals, measured in 1000 tonnes), and volume of transport (measured in million tonne-kilometres). Agricultural production variables are presented in Table 2 and classified across two dimensions as depicted in Table A1. The primary dimension pertains to whether a variable contributes by augmenting ('adding variable') or by depleting ('subtracting variable') the national inventory of cereals. 'Adding variables' augment the national inventory of cereals, while 'subtracting variables' encompass factors linked to consumption, utilisation, or sales. The secondary categorisation criterion is rooted in the variables' origin: whether they are derivative constructs resulting from other variables or distinct and unique entities. For instance, the change in stocks represents the disparity between closing and opening stocks, whereas closing and opening stocks themselves are distinct variables that stand independently.

Table 1. Transportation data of cereals [36].

Year	Vehicle Mileage, 1000 km [1]	Volume of Goods, 1000 tonnes [2]	Volume of Transport, Million tkms [3]
2011	15,895	2929	414
2012	7478	2146	217
2013	6248	2242	214
2014	7910	2340	271
2015	7816	2249	238
2016	11,300	2878	308
2017	10,477	2421	374
2018	6400	2183	184
2019	14,082	2574	520
2020	7314	2060	259
2021	10,178	2654	397
2022	9836	2307	298

[1] Total travelled journey of cereal-carrying trucks in 1000 km, [2] total carried amount of cereals in trucks in 1000 tonnes, [3] multiplication of vehicle mileage when trucks are loaded and amount of cereals in trucks.

Table 2. Agricultural data on cereals' production, stocks, and use.

Agri. Season	2009/2010	2010/2011	2011/2012	2012/2013	2013/2014	2014/2015	2015/2016	2016/2017	2017/2018	2018/2019	2019/2020	2020/2021	2021/2022
Opening stocks, 1.7.	1960	2832	1568	1623	1634	1634	1638	1714	1384	1396	1100	1352	1321
Production	4214	2942	3611	3605	3996	4057	3648	3527	3390	2691	3936	3321	2567
Import	52	50	52	73	80	56	33	27	50	68	16	11	195
Export	625	973	786	570	789	1010	663	780	501	401	799	566	335
Available, total	5601	4851	4444	4731	4663	4738	4656	4488	4324	3754	4253	4119	3748
Use as seed	259	275	241	247	286	276	266	254	256	252	260	260	289
Use for food, total	419	441	419	429	434	439	431	436	428	432	461	456	478
Use for food in industry	410	432	415	425	427	432	424	429	421	428	454	449	472
Use for food on farms	9	9	4	4	7	7	7	7	7	4	7	7	6
Use as feed, total	2066	2078	2236	1966	1916	2087	1925	2048	1886	1626	1772	1777	1730
Use as feed in industry	573	591	603	635	605	620	641	615	627	563	589	603	615
Use as feed on farms	1493	1487	1633	1330	1311	1467	1284	1433	1259	1063	1183	1174	1115
Industrial use	305	349	332	281	307	319	310	325	340	339	306	338	307
Domestic use, total	3054	3148	3235	2925	3008	3142	2953	3105	2925	2652	2821	2851	2811
Residual	−286	135	−413	−130	21	−42	−12	−2	4	2	80	−54	12
Closing stocks	2832	1568	1623	1936	1634	1638	1714	1384	1396	1100	1352	1321	925
Change in stocks	872	−1264	55	313	258	4	76	−330	11	−296	252	−31	−397
Use, residual, and closing stocks, total	5601	4851	4444	4731	4663	4738	4656	4488	4324	3754	4253	4119	3748
Use for energy on farms	5	7	6	2	65	21	21	42	14	3	22	20	7

Selection of Variables

Excluding 'residuals', there exists a total of 18 distinct potential independent variables available for inclusion in forecasting models associated with dependent variables. Indications suggest that some may exhibit correlation, potentially leading to multicollinearity. Multicollinearity significantly increases standard errors, leading to misleadingly insignificant statistical results [41]. Thus, it is important to avoid multicollinearity [42]. The selection of variables was guided by data availability. Initially, available variables in the balance sheet of cereals in Finland were included in the possible forecasting model, as they are relevant to road transport. These variables represent the production, utilisation, and stock levels of cereals, which are inherently tied to road transport. However, not all variables could be retained in the models. The exclusion of certain variables was dictated by the need to ensure statistical validity of the models, particularly with regard to multicollinearity. Retaining these variables could distort the results by inflating variance of regression coefficients.

To avoid multicollinearity between independent variables, covariance variables are revealed and excluded from a regression analysis through the variance inflation factor (VIF). VIF is calculated as shown in Equation (2). When VIF falls below 5, it denotes minimal covariance; between 5 and 10, a moderate level exists, while surpassing 10 signifies a substantial presence of multicollinearity [43]. The data covariance verification is usually satisfied by removing independent variables that cause covariance [34]. In this study, the threshold is taken as VIF < 5.

$$VIF_i = \frac{1}{1 - R^2}\ (i = 1,\ 2,\ 3\ldots) \tag{2}$$

The variables delineated in Table 3 below were omitted from subsequent analyses due to their VIF significantly exceeding the threshold of 5.

Some of the variables, such as 'available, total' and 'use, residual, and closing stocks, total', are closely related, with them often being components or subsets of others. That is why keeping such closely related variables together would introduce multicollinearity without adding new information. As a result, excluding such variables does not lead to any significant loss of information or affect the outcomes. The exclusion mainly serves to simplify the model and enhance its statistical validity.

In the context of absolute data, the variables omitted from the subsequent regression analysis owing to their high VIF in relation to other variables are 'available, total', 'use for food, total', 'use as feed, total', 'domestic use, total', and 'use, residual, and closing stocks, total'. In the context of annual percentage change, the variables of 'use, residual, and closing stocks, total', 'use for food, total', 'use as feed, total', and 'domestic use, total' were omitted from the analysis for the same reason. Besides the covariant variables, the variables 'import', 'use for energy on farms', and 'use for food on farms' were excluded from the analysis due to their notably low volumes.

Table 3. VIF of variables.

			Covariant Variables and VIF with Eliminated Variables				
			Closing stocks	Use, residual, and closing stocks, total	Use for food in industry	Use as feed on farms	Use as feed, total
Absolute data	Eliminated Variables	Available, total	>5.0	>5.0	<5.0	<5.0	<5.0
		Use for food, total	<5.0	<5.0	>5.0	<5.0	<5.0
		Use as feed, total	<5.0	<5.0	<5.0	>5.0	<5.0
		Domestic use, total	<5.0	<5.0	<5.0	>5.0	>5.0
		Use, residual, and closing stocks, total	>5.0	<5.0	<5.0	<5.0	<5.0
			Available, total	Use for food in industry	Use as feed on farms		
Annual percentage change	Eliminated Variables	Use, residual, and closing stocks, total	>5.0	<5.0	<5.0		
		Use for food, total	<5.0	>5.0	<5.0		
		Use as feed, total	<5.0	<5.0	>5.0		
		Domestic use, total	<5.0	<5.0	>5.0		

Owing to a limited number of observations, the models are restricted to accommodate only two independent variables. Potential independent variables that exhibited high VIF and those with notably lower volumes compared to other variables were excluded. Consequently, eleven prospective independent variables remain for the regression models that use absolute data. For the regression models using annual percentage change, there are twelve independent variables.

5. Results

With absolute data, there exist 55 potential models feasible for each dependent variable. When considering the annual percentage change and the inclusion of twelve prospective independent variables, the number of models can reach 66. Within Table 4, the adjusted R^2 values showcasing the relationship of each independent variable to the dependent variables are provided, alongside the corresponding p-values for the models. Additionally, Table 4 presents the optimal models, constrained to only two independent variables, delineating their combined adjusted R^2 and associated *p*-values Similarly, Table 5 presents performance and variables of 6 multiple regression models..

Table 4. Initial regression analysis: Adjusted R^2 and *p*-values for agricultural production variables.

Data Type	Dependent Variables	Adjusted R^2 and *p*-Value	Independent Variables Opening Stocks (T − 2/T − 1)	Production (T − 2/T − 1)	Production (T − 1/T)	Export (T − 2/T − 1)	Available, Total (T − 2/T − 1)	Use as Seed (T − 2/T − 1)	Use for Food in Industry (T − 2/T − 1)	Use as Feed in Industry (T − 2/T − 1)	Use as Feed on Farms (T − 2/T − 1)	Industrial Use (T − 2/T − 1)	Closing Stocks (T − 2/T − 1)	Change in Stocks (T − 2/T − 1)
Absolute data	Volume of goods (T)	Adjusted R^2	0.067	−0.072	0.514	−0.063	–	−0.072	−0.092	−0.095	−0.061	0.053	0.266	0.310
		p-Value	0.211	0.622	0.005	0.568	–	0.620	0.796	0.834	0.557	0.232	0.050	0.035
	Volume of transport (T)	Adjusted R^2	0.292	0.032	−0.011	0.040	–	−0.099	−0.098	−0.081	−0.003	−0.087	−0.040	0.139
		p-Value	0.041	0.270	0.370	0.256	–	0.941	0.886	0.686	0.349	0.733	0.464	0.126
	Vehicle mileage (T)	Adjusted R^2	0.336	−0.089	0.126	−0.056	–	−0.083	−0.057	−0.100	−0.098	−0.079	0.229	0.215
		p-Value	0.028	0.754	0.139	0.535	–	0.700	0.537	0.983	0.887	0.669	0.066	0.073
Annual percentage change	Volume of goods (T)	Adjusted R^2	0.303	0.259	0.497	−0.027	0.470	−0.106	−0.105	0.079	0.030	0.082	0.425	−0.095
		p-Value	0.046	0.063	0.009	0.412	0.012	0.847	0.827	0.206	0.283	0.202	0.018	0.728
	Vehicle mileage (T)	Adjusted R^2	0.141	0.222	−0.002	−0.078	0.059	−0.111	−0.067	0.105	−0.064	−0.049	0.213	−0.012
		p-Value	0.139	0.082	0.347	0.612	0.234	0.957	0.556	0.174	0.543	0.483	0.086	0.373
	Vehicle mileage (T)	Adjusted R^2	0.089	0.342	0.036	−0.097	0.204	−0.102	−0.087	0.108	−0.097	−0.058	0.296	−0.030
		p-Value	0.194	0.035	0.272	0.744	0.092	0.792	0.668	0.171	0.744	0.519	0.048	0.422

Table 5. Multiple regression analysis summary: Performance and variables of models.

Data Type		Model	Dependent Variable	Variables	Coefficients	95% Confidence Interval of Coefficients
Absolute Data	A [1]	Adjusted R^2: 0.6029 *p*-Value: 0.0064 RMSE [2]: 155.06	Volume of goods (T)	Production (T − 2/T − 1) Production (T − 1/T) Intercept	0.46 −0.11 1467	[0.19, 0.73] [−0.44, 0.05] [123, 2810]
	B	Adjusted R^2: 0.3841 *p*-Value: 0.0458 RMSE: 67.88	Volume of transport (T)	Production (T − 1/T) Change in stocks Intercept	−0.11 0.08 684	[−0.22, 0] [−0.03, 0.18] [300, 1069]
	C	Adjusted R^2: 0.5621 *p*-Value: 0.0099 RMSE: 1736. 47	Vehicle mileage (T)	Production (T − 1/T) Closing stocks (T − 2/T − 1) Intercept	−3.64 15,766 16,528	[−6.45, −0.83] [0.30, 6.56] [5162, 27,893]
Annual Percentage Change	D	Adjusted R^2: 0.6211 *p*-Value: 0.0084 RMSE: 0.10	Volume of goods (T)	Available, total (T − 2/T − 1) Opening stocks (T − 2/T − 1) Intercept	17,533 −0.36 0.027	[0.31, 2.65] [−0.74, 0.03] [−0.06, 0.11]
	E	Adjusted R^2: 0.2672 *p*-Value: 0.1181 RMSE: 0.49	Volume of transport (T)	Production (T − 1/T) Opening stocks (T − 2/T − 1) Intercept	−1.49 −1.02 0.11	[−3.64, 0.66] [−2.89, 0.86] [−0.28, 0.51]
	F	Adjusted R^2: 0.3839 *p*-Value: 0.059 RMSE: 0.34	Vehicle mileage (T)	Production (T − 1/T) Closing stocks (T − 2/T − 1) Intercept	−1.09 0.96 0.11	[−2.76, 0.58] [−0.78, 2.71] [−0.18, 0.39]

[1] Letters in this column represents the code of the models. [2] Root Mean Square Error.

Statistically, the data analysis is reliant on the number of the observations. For instance, the R-square value is notably affected by both the volume of data and the quantity of variables included. However, there is no guarantee that a substantial increase in sample size will proportionally augment the R^2 value of a regression model. This particularly applies when the method of data collection has been modified throughout the timeline or datasets are randomly drawn from surveys [44]. R^2 delineates the fraction of variance in the dependent variable explained by independent variables. Conversely, adjusted R^2 penalises the addition of excessive (and unnecessary) variables in the models. This provides a more reliable measure of the model's goodness of fit, particularly in scenarios involving multiple predictors. The desirable thresholds for R^2 and adjusted R^2 hinge significantly upon the discipline and complexity of models, lacking universally accepted standards. While adjusted R^2 provides a more robust model fit, it is important to also recognise its limitations. The relatively small number of observations in this study means that the adjusted R^2 might not fully capture more nuanced correlations between variables. To mitigate this, *p*-values were relied upon. However, even though risks are decreased by this way, they still exist. The results should therefore be interpreted with caution given the dataset's constraints. Despite this, production variables from previous years can potentially be integrated into road freight demand forecasting models.

One precondition established for the models in this study sets 0.55 as a threshold value for adjusted R^2. In addition to adjusted R^2 being >0.55, α is set as 0.05. This means that models wherein the *p*-value is <0.05 and adjusted R^2 is >0.55 are deemed acceptable, indicating a potential relationship between dependent and independent variables. The Type I error level (α) is often selected between the range of 0.05 and 0.1 [45,46]. Because the number of observations is limited in this study's dataset, a Type II error is not taken into account when checking so-called validity of the prospective models. Moreover, another important term in multiple linear regression models is the *confidence interval*, which is 95% in this study [47].

These specifications are met by three models, none of which pertain to the volume of transport. Three models show that agricultural production variables from the past have potential to explain variations in road freight demand. The independent variable of model A is the 'absolute value of volume of goods (y_1)', which may potentially be forecasted using the production amounts from the last two agricultural seasons: 'production (T − 2/T − 1) (x_1)' and 'production (T − 1/T) (x_2)' (See Figure 4).

$$y_1 = 0.46x_1 - 0.2x_2 + 1467 \tag{3}$$

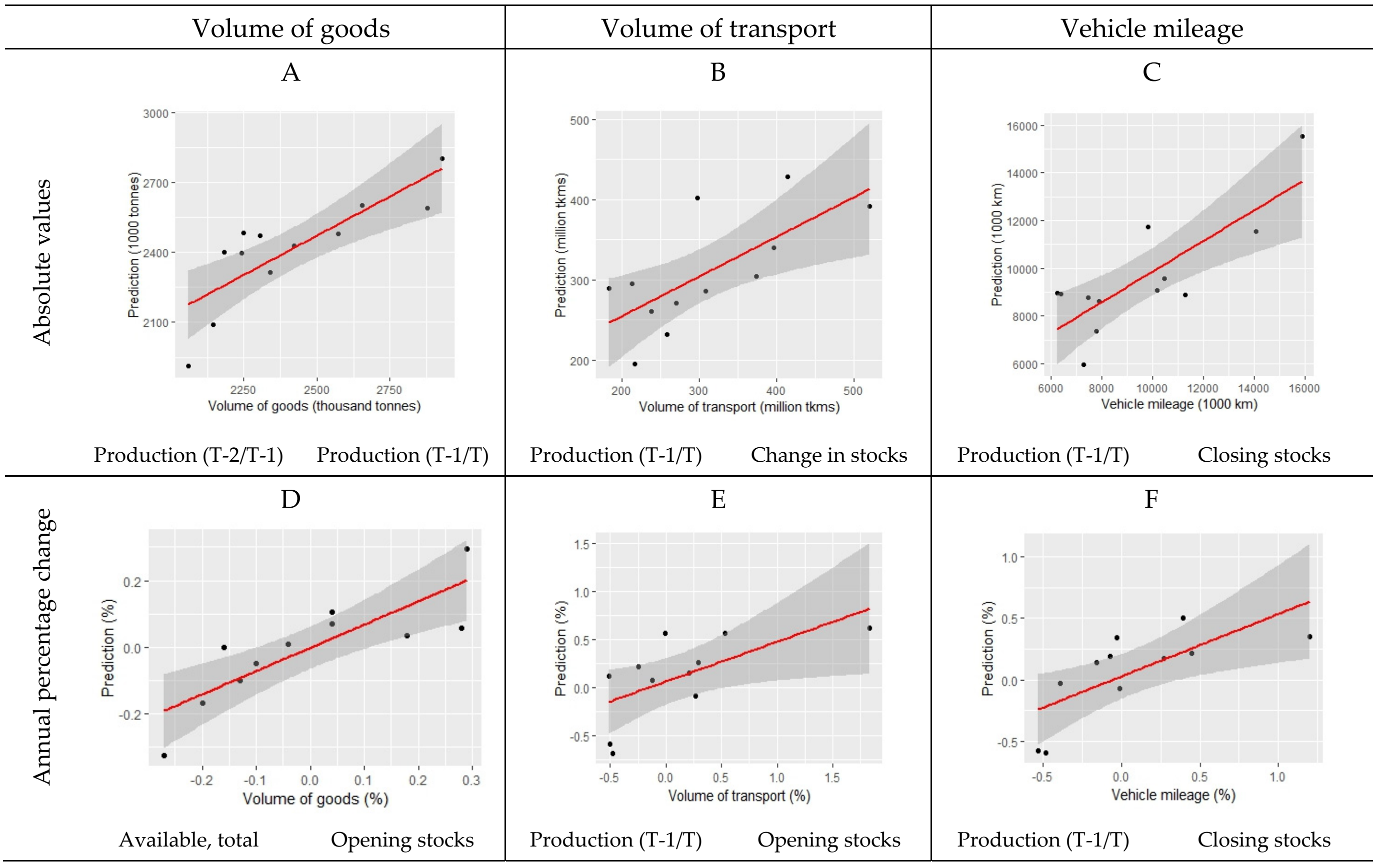

Figure 4. Best-fitting regression models (Red line represents the regression line, dark grey area represents confidence interval and black dots represent individual data points).

Model C pertains to 'absolute vehicle mileage (y_3)'. The predictive variables for this model are 'production (T − 1/T) (x_2)' and 'closing stocks (T − 2/T − 1) (x_5)' (See Figure 4).

$$y_3 = -3.64x_2 + 3.43x_5 + 16528 \tag{4}$$

Model D pertains to the 'annual percentage change in the volume of goods (y_2)'. The predictive variables for this model encompass 'available, total (T − 2/T − 1) (x_3)' and 'opening stocks (T − 2/T − 1) (x_4)' (See Figure 4).

$$y_2 = 1.48x_3 - 0.36x_4 + 0.027 \tag{5}$$

Overall, except these three models, the other models show lower adjusted R^2 values and higher *p*-values, indicating weaker fits. The overall results also suggest that past production data remain a valuable predictor in freight demand models.

6. Discussion

This study explores the correlation between historical agricultural variables and transport-related parameters, specifically, the volume of goods, transport volume, and vehicle mileage. Notably, the study reveals the predictability of the volume of goods and vehicle mileage through current statistical information, contrasting with the less predictable nature of transport volume. Emergent findings suggest that 60% of the absolute value of the volume of goods (transported cereals) can potentially be forecasted based on the production from the previous years and two years prior (model A). Similarly, 62% of the annual percentage change in the volume of goods (cereals) may be estimated through the available, total, and opening stocks (model D). Additionally, 56% of the absolute value of vehicle mileage can potentially be predicted using the last year's production and closing stocks from two years prior (model C) (see Figure 5).

An examination of the potential correlation between historical agricultural production variables and transport-related parameters appears to be rather absent in the existing literature. Consequently, this investigation serves to address this gap, characterising itself with an exploratory nature, seeking to uncover and comprehend the nuances of such a relationship. This study employs a multiple linear regression analysis to investigate the potential relationship. However, it is worth noting that there is no consensus on the minimum number of observations for multiple linear regression analyses, with varying recommendations across disciplines [48,49]. This study is constrained by a limited number of observations, viz. 12 for absolute data and 11 for the annual percentage change analysis. While insufficient for definitive conclusions, these observations offer invaluable insights into the potential relationships with certain variables. Notably, despite the availability of annual data since 1995, the analysis commences from 2009 due to variations in data collection methods before this period, yielding inconsistent data inputs.

When considering why certain variables are in the models but not others, a few prospective reasons come to light. First, the consistency and reliability of independent variables in the valid models might surpass others. Second, these key variables are perhaps more directly linked to how goods move through the agricultural supply chain. Third, their impact on the agricultural supply chain may manifest more promptly, in contrast to other variables that exert delayed and indirect influences on transport-related factors.

In model A (where the volume of goods amount is explained best by production from the last two years), an intriguing observation emerges. The negative sign for the last year's production suggests a temporal lag in its impact on the volume of goods. This negative sign may also be linked to storage dynamics. If goods from the last year are stored and carried over to the current period, the need for immediate transportation might be reduced. Furthermore, differing signs for production variables could imply that a surplus from the last season prompts a scaling back of transportation efforts to optimise costs.

When interpreting model C, it appears that increased closing stocks from one agricultural season ($T - 2/T - 1$) to opening stocks of the next ($T - 1/T$) seem to positively influence vehicle mileage in year T. This could imply that higher end-of-season stocks may require more transport efforts, possibly for managing or distributing excess goods. Additionally, the model suggests a negative correlation between production in season $T - 1/T$ and vehicle mileage, proposing a delicate balance where increased production could lead to less vehicle mileage. This might reflect more upon efficient production and shipping strategies. However, this link between production efficiency and transportation logistics should be considered tentatively, recognising the potential oversimplification of supply chain realities.

Model D reveals how inventory levels influence the annual percentage change in transported goods. This suggests that instead of production, it is the inventory levels that are closely linked to the changes in the volume of goods transported. The negative sign for opening stocks in the model might mean that having more cereals in stock leads to less immediate need for transportation. The time gap between the variables is also considerably important here. Even though 'available, total' technically includes opening stocks, the model hints that these two factors together play a role. They help in explaining some of the perturbations observed in how much cereals get moved around each year.

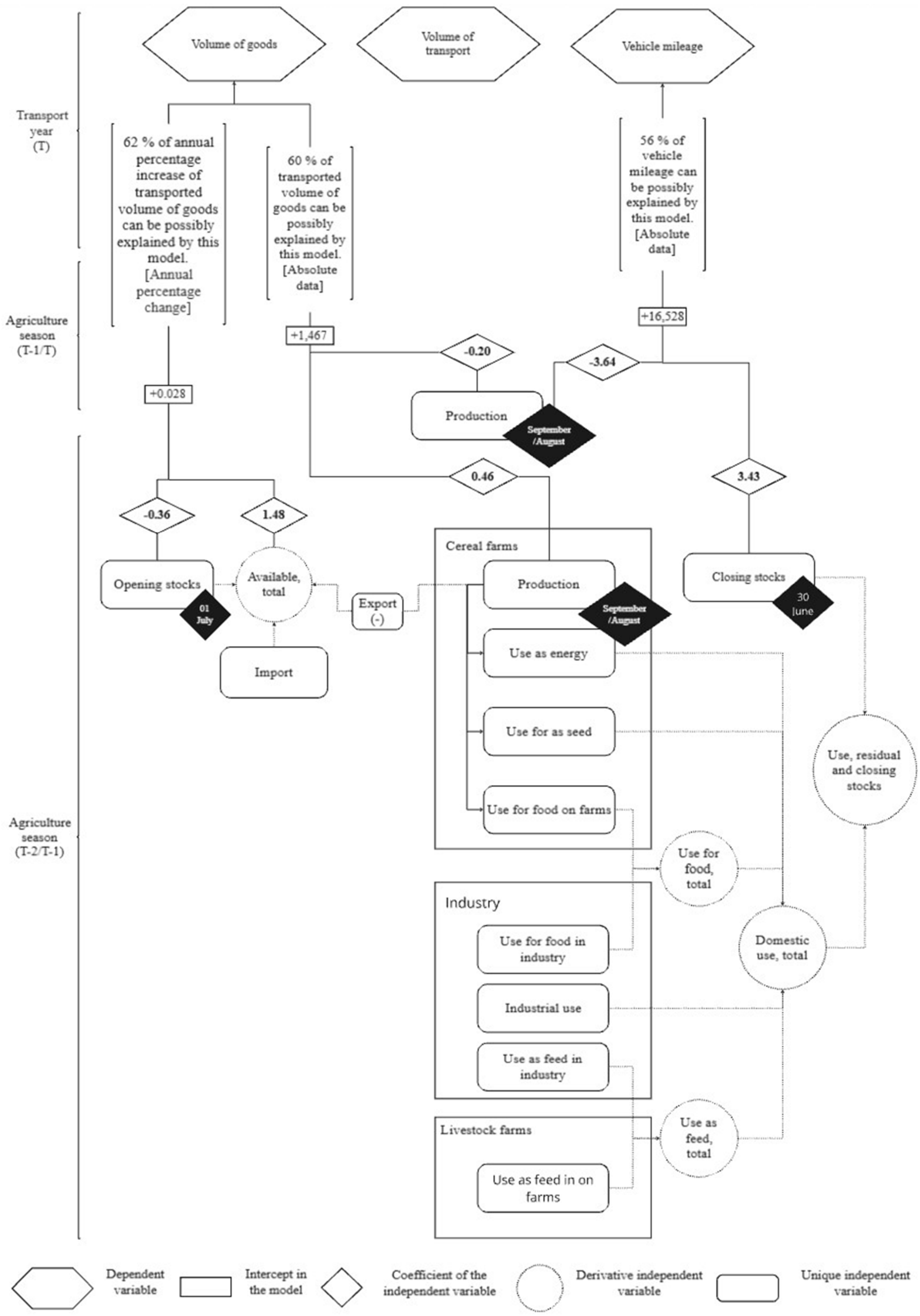

Figure 5. Relationship of variables.

6.1. Theoretical and Practical Implications

This study contributes to the broader dialogue on road freight demand forecasting by exploring the relationship between agricultural production data and road freight-related parameters, with a focus on cereals. It suggests that incorporating specific agricultural production variables, such as production and inventory levels, could enhance the accuracy of existing models. On a practical level, the insights gained from this study may assist in the incremental improvement of transport planning and logistics management within the agriculture industry.

While national freight forecasting models like Samgods, the Norwegian Freight Transport Model, the TRIMODE model, and FRISBEE provide frameworks for understanding overall transport demand, these models do not account for specific requirements of agricultural transport [12–14,24]. Instead, these models predominantly focus on aggregate flows and logistical optimisation at the national or European level, leaving a gap when it comes to integrating production-specific data such as yearly agricultural output and inventory levels into the forecasting process. This research offers a granular view of transport demand that existing models do not currently provide. This approach underscores the importance of sector-specific variables in improving logistical efficiency of freight systems. Consequently, this study complements existing models by introducing a perspective that emphasises the interaction between production cycles and transportation logistics in agriculture. However, there is no doubt that similar cycles could be found from other industry sectors, and certainly regarding other commodities within agriculture.

Recognising the potential impact of agricultural production and stock levels on transport demand could help in fine-tuning resource allocation and logistics strategies. This study further encourages a careful consideration of how localised and sector-specific data might be leveraged to optimise transportation efficiency, although acknowledging the complexity and variability of transport dynamics across different regions and agricultural practises. The regression models presented in this study demonstrate that production and inventory levels from prior years have potential for serving as predictors for road freight demand, particularly for the volume of goods transported. These models offer a practical framework for integrating agricultural production data into transport forecasting systems. The beneficiaries of more precise road freight demand derive from several directions. The first and the most obvious one is the logistics industry, especially the road freight operators specialising in agricultural products. Investments made in their fleet and the business planning they undertake regarding their customer segments are dependent on somewhat credible demand forecasting abilities. Second, there is another market to be considered. The original manufacturers of the vehicles and specialised equipment can more reliably make projections on how the fleet demand is developing, i.e., how their sales are expected to grow or decline. Finally, there is a general interest from the infrastructure policy point of view. Road freight vehicles stress infrastructure to a considerable extent and therefore any changes in the freight volumes will have implications on infrastructure asset management and lifecycles. Being able to forecast, even roughly, the projections for road freight volumes in terms of time and place will offer possibilities to optimise asset management strategies of the ministries and central agencies responsible for road infrastructures.

6.2. Limitations and Recommendations for Future Research

This approach invites further investigation into how detailed sector-specific data can be systematically integrated into transport forecasting methodologies, potentially offering a more detailed understanding of the factors influencing road freight demand. However, the study encounters limitations that must be acknowledged. First, statistical significance of the observed relationships is constrained by the relatively short span of time series data available. This limitation inherently impacts the robustness of our predictive models and may mask more profound correlations that could emerge from a longer time series. Second, our analysis primarily relies on publicly available data, which may not capture all the details of agricultural production and freight dynamics. The study's focus on cereals as a case product group, while providing more in-depth insights, limits generalisability of the findings across different agricultural commodities. To address these limitations and build upon the foundational work of this study, several avenues for future research are proposed. As more time series data become available, future studies can build upon these seminal research findings, improving the statistical validity and identifying trends and relationships over time. Extending the investigation to include different agricultural commodities will reveal whether the observed patterns hold across various sectors, each with unique logistics and road freight requirements. A geographic analysis is also recommended to examine regional differences in production and road freight. Furthermore, future studies could consider using longitudinal or panel datasets to apply fixed effects modelling, allowing for a deeper exploration of causal relationships and more comprehensive insights into the influence of agricultural production variables on transport demand. This suggested path forward will deepen the understanding of the relationship between agriculture and road freight, contributing to more effective and sustainable logistics planning as new data become available.

7. Conclusions

This study explored the relationship between historical agricultural variables and key road freight parameters, viz. the volume of freight in tonnes, freight volumes in tonne-kilometres, and vehicle kilometres. Findings highlight that the production variables do have some predictive power for road freight demand, albeit the predictive power is not significant based on the rather short time series data that were available for this research. However, the results call for more research in the future as the dataset for the time series increases in volume. The absolute volume of goods (cereals) can potentially be forecasted based on the production from the previous years and two years prior, while annual percentage change may be estimated through available, total, and opening stocks. Similarly, vehicle mileage can potentially be predicted using the last year's production and closing stocks from two years prior. This result has relevance when agricultural production volumes are forecasted in the longer term, since it seems to have implications for expected road freight demand.

To integrate agricultural production data into national forecasting models, first, national statistical agencies should focus on collecting and harmonising agricultural production and inventory data, and ensure consistency across years and sectors. This would make it easier to include the data in broader transport models. Second, incorporating sector-specific variables like production and stock levels into existing national freight forecasting models is suggested. Lastly, collaboration between agriculture, logistics, and transport policymakers is encouraged to integrate these data into practical forecasting applications.

This study addresses a gap in the existing literature, being the first to explore the correlation between historical agricultural production volumes and road freight demand. By this means, this study suggests an approach for more efficient logistical planning, which is essential for optimising resource allocation. Integrating agricultural production data into freight forecasting models has the potential to improve forecasting accuracy and support the development of sustainable agricultural logistics. In doing so, this research contributes to the broader literature on sustainable logistics in agriculture by highlighting how production cycles and inventory management can be leveraged to enhance transport efficiency and reduce environmental impact. Through the non-complex application of a multiple linear regression analysis, the study offers a perspective on the complicated relationships that are worthy of deeper exploration. While acknowledging the constraints due to a limited number of observations, the basic idea seems to bear relevance in the light of the results, thus providing a foundation for more sophisticated road freight demand modelling.

The presence of specific variables in the models underscores their consistency, reliability, and direct relevance to the agricultural supply chain. The negative signs in the models are most likely associated with a temporal lag. Furthermore, storage dynamics influence the volume of goods transported. This study contributes practically by enhancing our understanding of agriculture–transport dynamics and provides a theoretical foundation for further exploration in the evolving field of agricultural transport forecasting.

Author Contributions: Conceptualisation, T.K. and P.L.; methodology, T.K. and P.L.; software, T.K.; validation, P.L., D.J.E. and T.K.; formal analysis, T.K. and P.L.; investigation, P.L. and D.J.E.; resources, T.K.; data curation, T.K.; writing—original draft preparation, T.K.; writing—review and editing, P.L.,

D.J.E. and T.K.; visualisation, T.K.; supervision, P.L. and D.J.E.; project administration, P.L. and T.K.; funding acquisition, P.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Business Finland, grant number 3509/31/2021.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: This study uses publicly available data from the National Statistical Institute of Finland (https://statfin.stat.fi/PxWeb/pxweb/en/StatFin_Passiivi/StatFin_Passiivi__kttav/statfinpas_kttav_pxt_909_201000_fi.px/ (accessed on 15 August 2024)) and from National Resources Institute Finland (https://statdb.luke.fi/PxWeb/pxweb/en/LUKE/LUKE__02%20Maatalous__04%20Tuotanto__32%20Viljatase/01_Viljatase.px/?rxid=dc711a9e-de6d-454b-82c2-74ff79a3a5e0 (accessed on 15 August 2024)).

Acknowledgments: We kindly thank LEVITOI research project members and statisticians in both the National Statistical Institute of Finland and National Resources Institute Finland for their guidance and support in the process.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Explanation of agricultural production variables.

Variable	Subtracting/ Adding	Unique/ Grouping Variable	Explanation
Production	Adding	Unique	Produced cereals within agricultural season
Import	Adding	Unique	Imported cereals within agricultural season
Opening stocks as of 1.7.	Adding	Unique	Amount of stocks in Finland at beginning of agricultural season
Available, total	Adding	Derivative	(Production + import + opening stocks as of 1.7) − export
Export	Subtracting	Unique	Exported cereals within agricultural season
Closing stocks, 30.6.	Subtracting	Unique	Amount of stocks in Finland at end of agricultural season
Use as feed on farms	Subtracting	Unique	Used cereals as feed on livestock farms
Use for food on farms	Subtracting	Unique	Used cereals for farmers' own consumption
Use as seed	Subtracting	Unique	Used cereals for seeding by farmers
Use for energy on farms	Subtracting	Unique	Used cereals burned for energy on farms
Use as feed in industry	Subtracting	Unique	Used cereals in industry for sending back to livestock farms for feed use
Use for food in industry	Subtracting	Unique	Used cereals as processing and packaging by industry for human consumption
Industrial use	Subtracting	Derivative	Nonfood and nonfeed industrial production
Use as feed, total	Subtracting	Derivative	Use as feed on farms + use as feed in industry
Use for food, total	Subtracting	Derivative	Use for food on farms + use for food in industry
Domestic use, total	Subtracting	Derivative	Use for food on farms + use as seed + use as energy on farms + use as feed on farms + use as feed in industry + use for food in industry
Use, residual, and closing stocks, total	Subtracting	Derivative	Domestic use, total + closing stocks + residual
Change in stocks	Unclassified	Derivative	Closing stocks, 30.06 − opening stocks as of 1.7
Residual	Unclassified	Unclassified	Unexplained in statistics

References

1. Cuéllar, A.D.; Webber, M.E. Wasted Food, Wasted Energy: The Embedded Energy in Food Waste in the United States. *Environ. Sci. Technol.* **2010**, *44*, 6464–6469. [CrossRef] [PubMed]
2. Porter, S.D.; Reay, D.S. Addressing Food Supply Chain and Consumption Inefficiencies: Potential for Climate Change Mitigation. *Reg. Environ. Chang.* **2016**, *16*, 2279–2290. [CrossRef]
3. Péra, T.G.; da Rocha, F.V.; Caixeta Filho, J.V. Tracking Food Supply Chain Postharvest Losses on a Global Scale: The Development of the Postharvest Loss Information System. *Agriculture* **2023**, *13*, 1990. [CrossRef]
4. Eurostat. *Agriculture, Forestry and Fishery Statistics—2020 Edition*; Eurostat: Luxembourg, 2020; ISBN 9789276215219.
5. Karasu, T.; Hussain, S.; Leviäkangas, P. First Mile Challenges for Agricultural Logistics. In *Lecture Notes in Logistics*; Clausen, U., Dellbrüge, M., Eds.; Springer: Cham, Switzerland, 2023; pp. 60–74. [CrossRef]
6. Miklius, W.; Garrod, P.V.; Casavant, K.L. Estimation of Demand for Transportation of Agricultural Commodities. *Am. J. Agric. Econ.* **1976**, *58*, 217–223. [CrossRef]
7. Fazekaš, T.; Bobera, D.; Ćirić, Z. Ecologically and Economically Sustainable Agricultural Transportation Based on Advanced Information Technologies. *Econ. Agric.* **2017**, *64*, 739–751. [CrossRef]
8. Zhang, D.; Sun, Z. The Impact of Agricultural Global Value Chain Participation on Agricultural Total Factor Productivity. *Agriculture* **2023**, *13*, 2151. [CrossRef]
9. Zhang, C.; Chen, D. Fragmentation Reduction through Farmer-Led Land Transfer and Consolidation? Experiences of Rice Farmers in Wuhan Metropolitan Area, China. *Agriculture* **2021**, *11*, 631. [CrossRef]
10. Sun, S.; Gu, M.; Ou, J.; Li, Z.; Luan, S. Regional Truck Travel Characteristics Analysis and Freight Volume Estimation: Support for the Sustainable Development of Freight. *Sustainability* **2024**, *16*, 6317. [CrossRef]
11. Pastinen, V.; Salanne, I.; Keränen, M.; Lehto, H.; Jaakkola, E.; Tikkanen, M. National Model System for Transport—A Study of the Prerequisites and Options for Developing a Model System. Available online: https://www.traficom.fi/sites/default/files/media/publication/Liikenteen%20valtakunnallinen%20mallij%C3%A4rjestelm%C3%A4.pdf (accessed on 8 May 2024).
12. Fiorello, D.; Nökel, K.; Martino, A. The TRIMODE Integrated Model for Europe. *Transp. Res. Procedia* **2017**, *31*, 88–98. [CrossRef]
13. Sala, G.; Romano, D.; Berglund, M. Samgods User Manual V1.0. Available online: https://bransch.trafikverket.se/contentassets/d7cf7d727fb2488aab9fa9d24387c7c8/externa-rapporter/samgods_user_manual_v1.0_rad81.pdf (accessed on 12 February 2024).
14. Madslien, A.; Steinsland, C.; Gronland, S.E. *The National Norwegian Freight Transport Model. How to Use the Model*; The Institute of Transport Economics: Oslo, Norway, 2015.
15. Gaudry, M.J.I. External Review of the Danish Landstrafikmodel LTM 3.0 Project. 2017. Available online: https://www.researchgate.net/publication/323317468 (accessed on 8 May 2024).
16. Petrik, O.; De Abreu e Silva, J.; Moura, F. Stated Preference Surveys in Transport Demand Modeling: Disengagement of Respondents. *Transp. Lett.* **2016**, *8*, 13–25. [CrossRef]
17. Jayanthi, G.; Jothilakshmi, P. Traffic Time Series Forecasting on Highways—A Contemporary Survey of Models, Methods and Techniques. *Int. J. Logist. Syst. Manag.* **2021**, *39*, 77–110. [CrossRef]
18. Bilbao-Ubillos, J.; Fernandez-Sainz, A.; Heidenreich, N.; Sperlich, S. Flexible Estimation of Transport Demand Functions: Recommendations for Public Policy Makers. *Transp. Lett.* **2015**, *7*, 241–251. [CrossRef]
19. Norouzian-Maleki, P.; Izadbakhsh, H.; Saberi, M.; Hussain, O.; Jahangoshai Rezaee, M.; GhanbarTehrani, N. An Integrated Approach to System Dynamics and Data Envelopment Analysis for Determining Efficient Policies and Forecasting Travel Demand in an Urban Transport System. *Transp. Lett.* **2022**, *14*, 157–173. [CrossRef]
20. Garthwaite, I.; Raymond, W. Transporting Agricultural Goods. 2015. Available online: https://thedocs.worldbank.org/en/doc/538601534212778854-0050022018/original/EBA15Transportingagriculturegoods.pdf (accessed on 17 January 2024).
21. Finnish Cereal Committee. Production of Cereal and Oilseed Crops in Finland. 2014. Available online: https://www.vyr.fi/ (accessed on 17 January 2024).
22. Luke. Finland's Cereal Production Forms the Basis of Domestic Food—As Cereal, Dairy and Meat Products. 2023. Available online: https://www.luke.fi/en/news/ (accessed on 17 January 2024).
23. Kii, M.; Nakanishi, H.; Nakamura, K.; Doi, K. Transportation and Spatial Development: An Overview and a Future Direction. *Transp. Policy* **2016**, *49*, 148–158. [CrossRef]
24. Lahtinen, H.; Pulli, J. Developer's Handbook from Logistics Centre, Developing a Logistics Centre System for Southern Finland—Project 2009–2012. 2012. Available online: http://www.eslogc.fi/images/stories/ESLogC_kasikirja_web.pdf (accessed on 22 March 2024). (In Finnish).
25. Salanne, I.; Mäkelä, K.; Tikkanen, M. MERIMA—Finland's International Maritime Transport Emissions Models: Performance Report 2005–2019. 2021. Available online: https://www.traficom.fi/sites/default/files/media/publication/MERIMA_Tulosraportti_2005-2019_26012021_FINAL.pdf (accessed on 8 May 2024).
26. Salanne, I.; Tikkanen, M.; Saarto, P.; Leskinen, T. TRAMA—The Economical Effects of Transit. In *Transport–Results Report 2011*; Sito Group: Helsinki, Finland, 2011.
27. Moilanen, P.; Lapp, T.; Niinikoski, M.; Blomqvist, P.; Rinta-Piirto, J. *National Transport Forecasts*; Finnish Transport and Communications Agency Traficom: Helsinki, Finland, 2022.
28. Edwards, D.; Parn, E.A.; Sing, M.C.P.; Thwala, W.D. Risk of Excavators Overturning: Determining Horizontal Centrifugal Force When Slewing Freely Suspended Loads. *Eng. Constr. Archit. Manag.* **2019**, *26*, 479–498. [CrossRef]

29. Popper, K.R. *Objective Knowledge: An Evolutionary Approach*; Oxford University Press: Oxford, UK, 1972.
30. Halfpenny, P. *Positivism and Sociology: Explaining Social Life*; Routledge Taylor & Francis Group: London, UK, 1982; ISBN 9781138995161.
31. Edwards, D.J.; Rillie, I.; Chileshe, N.; Lai, J.; Hosseini, M.R.; Thwala, W.D. A Field Survey of Hand—Arm Vibration Exposure in the UK Utilities Sector. *Eng. Constr. Archit. Manag.* **2020**, *27*, 2179–2198. [CrossRef]
32. Luke. 'About'. Available online: https://www.luke.fi/en/about-luke (accessed on 8 December 2023).
33. Tilastokeskus. 'About Us'. Available online: https://stat.fi/org/index_en.html (accessed on 15 November 2023).
34. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*, 3rd ed.; OTexts: Melbourne, Australia, 2021.
35. Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis: Forecasting and Control*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
36. Tilastokeskus. 119b—Volume of Goods, Vehicle Mileage and Transport Performance of Domestic Road Transport by Type of Goods, 2011–2023. Available online: https://statfin.stat.fi/PxWeb/pxweb/en/StatFin_Passiivi/StatFin_Passiivi__kttav/statfinpas_kttav_pxt_909_201000_fi.px/ (accessed on 15 November 2023).
37. Li, C.; Liu, W. Multimodal Transport Demand Forecasting via Federated Learning. *IEEE Trans. Intell. Transp. Syst.* **2023**, *25*, 4009–4020. [CrossRef]
38. Gerardo Muros Anguita, J.; Díaz Olariaga, O. Air Cargo Transport Demand Forecasting Using ConvLSTM2D, an Artificial Neural Network Architecture Approach. *Case Stud. Transp. Policy* **2023**, *12*, 101009. [CrossRef]
39. Chen, B.; Wu, J. Predicting Model for Air Transport Demand under Uncertainties Based on Particle Filter. *Sustainability* **2022**, *14*, 16694. [CrossRef]
40. Sun, S.C.; Duan, Z.Y.; Chen, C. Freight Transport Demand Forecasting in Urban Logistics Planning: A Case Study of Yiwu City. *Appl. Mech. Mater.* **2014**, *505–506*, 915–921. [CrossRef]
41. Draper, N.R.; Smith, H. *Applied Regression Analysis*, 3rd ed.; Draper, N.R., Smith, H., Eds.; Wiley: Hoboken, NJ, USA, 1998.
42. Han, Z.; Youn, Y.C.; Kim, S.; Choe, H. Improving Farmer Livelihood Resilience to Climate Change in Rural Areas of Inner Mongolia, China. *Agriculture* **2023**, *13*, 2030. [CrossRef]
43. O'Brien, R.M. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Qual. Quant.* **2007**, *41*, 673–690. [CrossRef]
44. Rasyidah; Efendi, R.; Nawi, N.M.; Deris, M.M.; Burney, S.M.A. Cleansing of Inconsistent Sample in Linear Regression Model Based on Rough Sets Theory. *Syst. Soft Comput.* **2023**, *5*, 200046. [CrossRef]
45. European Medicine Agency. ICH Topic E 9, Statistical Principal for Clinical Trials Step 5, ICH Harmonised Tripartite Guideline. Available online: https://www.ema.europa.eu/ (accessed on 12 February 2024).
46. Chow, S.-C.; Shao, J.; Wang, H.; Lokhnygina, Y. *Sample Size Calculation in Clinical Research*, 3rd ed.; Chapman and Hall/CRC: New York, NY, USA, 2017.
47. Serdar, C.C.; Cihan, M.; Yücel, D.; Serdar, M.A. Sample Size, Power and Effect Size Revisited: Simplified and Practical Approachin Pre-Clinical, Clinical and Laboratory Studies. *Biochem. Med.* **2021**, *31*, 27–53. [CrossRef] [PubMed]
48. Rodríguez del Águila, M.M.; Benítez-Parejo, N. Simple Linear and Multivariate Regression Models. *Allergol. Immunopathol.* **2011**, *39*, 159–173. [CrossRef]
49. Austin, P.C.; Steyerberg, E.W. The Number of Subjects per Variable Required in Linear Regression Analyses. *J. Clin. Epidemiol.* **2015**, *68*, 627–636. [CrossRef]