Carbon footprint estimation and data sampling method: a case study of ecologically cultivated rice produced in Japan

Naoki Yoshikawa[1], Tomohiro Ikeda[1], Koji Amano[1], Koji Shimada[2]

[1] College of Science and Engineering, Ritsumeikan University,
1-1-1 Nojihigashi, Kusatsu, Shiga, 525-8577, Japan,
[2] College of Economics, Ritsumeikan University,
1-1-1 Nojihigashi, Kusatsu, Shiga, 525-8577, Japan[3]

ABSTRACT

Carbon footprint estimation of food products is considered to require collecting data on a number of agricultural producers to ensure statistical representativeness of inventory data. This study evaluated the carbon footprint of ecologically cultivated rice produced in Japan and examined the representativeness of inventory data employing survey sampling theory. Five life cycle stages were set for estimation: raw-material production, rice polishing, distribution and retailing, rice cooking, and waste treatment. Foreground data on over 100 producers were collected in agricultural production. The results show that the carbon footprint of rice is 7.7 kg-CO_2eq/package (4 kg of polished rice). The contribution of raw-material production is considerable, especially that of methane emissions from paddy fields. Representativeness is examined by the standard-error ratio of estimated inputs. The standard error ratio of greenhouse gas (GHG) emissions evaluated by poststratified estimator was 3.8%, which seemed to have enough representativeness. However, the results suggested a smaller sample can improve representativeness if implementing an optimal sample survey.

Keywords: carbon footprint, rice, data sampling

1. Introduction

Japanese activities related to the carbon footprint of products (CFP) started in 2008, and have reached the stage of sale in stores. Regarding carbon footprint estimation of food products, although there still is no consensus on data collection based on statistical theory, researchers may have to survey foreground data on a number of agricultural producers to ensure representativeness of inventory data. This might make CFP in food and agriculture unaffordable, especially for smaller suppliers, or unreliable without reasonable guidelines for data collection on mass suppliers. This study estimated the carbon footprint of ecologically cultivated rice produced in Shiga prefecture, Japan, which is the first product sold in stores to carry a carbon footprint label. In addition, we examined the representativeness of inventory data and the data collection methods, utilizing survey-sampling theory.

2. Estimating the carbon footprint of rice

2.1. Summary of CFP calculation

The product subject to estimation of CFP is specially cultivated polished rice (variety: Koshihikari) produced in the northern area of Shiga Prefecture, Japan (Figure 1). This product is treated with less than one-half the conventional application of chemical nitrogen fertilizer and agrochemicals in rice cultivation. Beginning in January 2010, packages with a CFP label have been sold in retailers around Japan. The functional unit in this study is one package (4 kg polished rice). GHG (CO_2, CH_4, and N_2O) emissions were estimated employing a cradle-to-grave analysis.

2.2. System boundaries

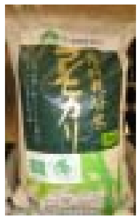

Figure 1: Product subject to CFP estimation

Five life cycle stages of rice were set for estimation: raw material production, rice polishing, distribution and retailing, rice cooking, and waste treatment. Figure 2 shows the system boundary of each stage.

In rice polishing stage, both the main product (polished rice) and co-products (rice bran, utilized as fertilizer material) are produced.The environmental loads of both products in the rice-cultivation and rice polishing stages were allocated by economical value.

Environmental load related to durables (agricultural equipment, facilities, cooking equipment, etc.) are not included because of uncertainty about their durable periods. Waste-recycling processes are not estimated in order to avoid double counting with utilization of recycled materials. Transportations of consumers between their homes and

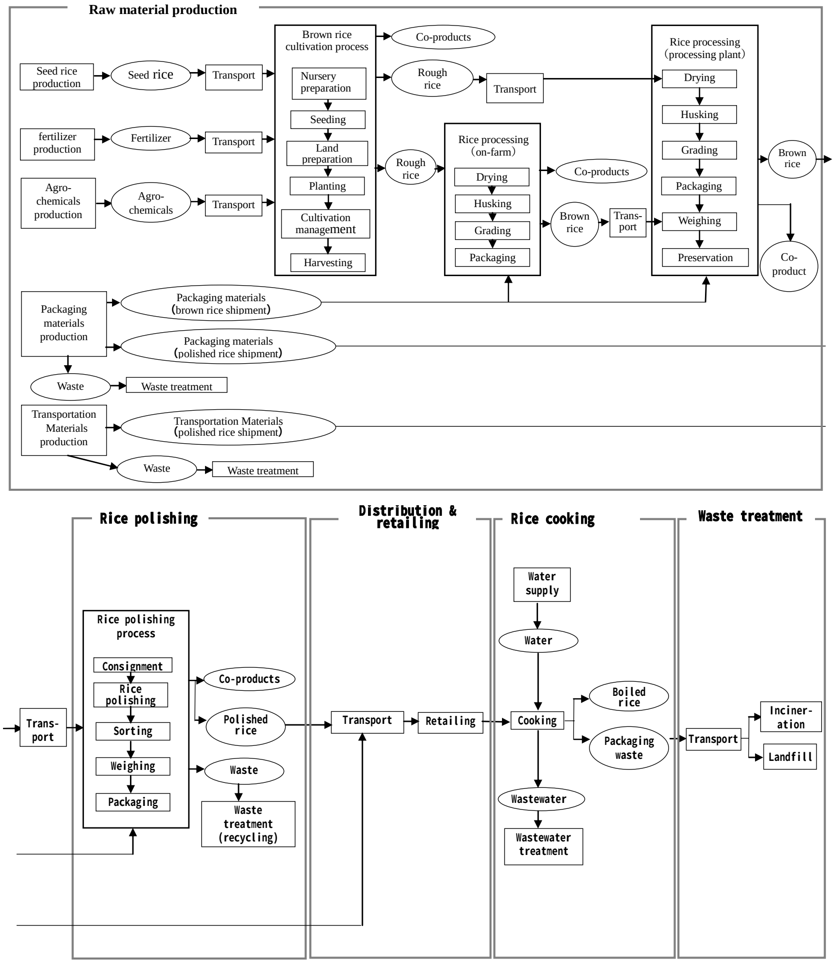

Figure 2: System boundaries

retailers are also not taken into account.

2.3. Data collection

Activity data were collected as foreground data when possible, though some data were collected as background data. Major input materials in each stage are summarized in Table 1.

In the raw material production stage, over 400 producers cultivate the rice for the subject product. This study collected data on 109 producers. These data cover over 50% of all the products, which the current Japanese carbon footprint calculation rules (Product Category Rules, or PCR) for rice require as the standard for data collection. Input data of fertilizer, agrochemicals, fuels, and electricity, in each agricultural producer and rice-processing plant, were surveyed. CH_4 and N_2O emissions from paddy fields also were taken into consideration (GIO, 2009). Actual data of transportation distance were collected for the main product; the distance (500 km) and loading factor scenarios were used for transport of inputs.

Foreground data were surveyed in the rice polishing stage and the distribution and retailing stage. Emissions from the rice polishing stage were calculated from energy usage in rice-polishing plants. Energy use in retailers was collected from chain stores dealing in the subject product. Data on whole stores were allocated to each product by calculating the emission factor per retail price. The average transport distance between stores and rice polishing plants was used for transport of packaged products based on past records of delivery. In the cooking stage, we utilized the PCR scenario, which includes average electricity and water use data in rice cooking using an average domestic rice cooker. In the waste treatment stage, we estimated data for incineration and disposal in landfills of plastic rice packages. The ratio of treatments used is the average value in Japan.

Table 1: Summary of data collection

Life cycle stage	*Inputs*	*Data source of background data*	*Life cycle stage*	*Inputs*	*Data source of background data*
Raw material production	Energy	JEMAI, 2009a	Rice polishing	Energy	JEMAI, 2009a
	Fertilizer	JEMAI, 2009b	Distribution & retailing	Energy	
	Agrochemicals			Transportation	
	Packaging materials	JEMAI, 2009a	Cooking	Energy	
	Seeds	Ajinomoto Co., Inc.,2007		Water supply	
	GHG from paddy field	GIO, 2009	Waste treatment	Waste treatment	

2.4. Results of carbon footprint estimation

Figure 3 shows the results of carbon footprint estimation per package (4 kg polished rice). CFP in all stages is 7.7 kg-CO_2eq/package. About 65% of emissions were related to the raw material production stage; almost all emissions come from agricultural production. CH_4 emission from paddy fields, which is caused by anaerobic fermentation, accounts for 50% of LC-GHGs from agricultural production, although uncertainty concerning its emission factor is high. Besides CH_4 emission, emission of GHGs from fertilizer, energy, and transportation of input materials each accounted for more than 5% of LC-GHGs in agricultural production. After the raw material production stage, the distribution and retailing stage and the rice cooking stage are key stages for emission of LC-GHGs. Most emissions in the cooking stage were

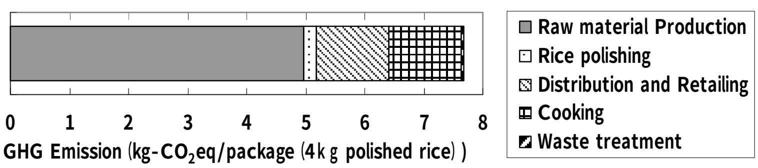

Figure 3: Results of carbon-footprint calculation

from electricity used by rice cookers. All transportation of products and inputs accounted for 6.5% of LC-GHGs.

3. Evaluation of representativeness of inventory data

3.1. Approach

When calculating the CFP of agricultural production using activity data surveyed by data sampling, uncertainty related to statistical errors in process data becomes an issue, as well as uncertainty of emissions factors and system boundaries. If implementing inadequate data sampling, the cost of surveying CFP rises to ensure reliability of activity data. Sampling survey theory can be applicable for evaluating agricultural activities involving a large number of small producers. This study evaluated the representativeness of CFP data by estimating the variability of calculated data and considered optimal data sampling.

Data variability is examined by the standard error ratio of material input quantity by parent population (all producers), estimated from data on sampled producers. Standard error ratio, corresponding to coefficient of variance of estimates, is evaluated by uncertainty of input data and sampling ratio from parent population. This indicates representativeness of inventory data because both average inventory data estimated from data with high uncertainties and that from few samples have poor reliability to use the data as representative data.

Since cultivated area varies by producer as seen in Table 2, it is assumed that the input quantity of each material correlated with cultivation area. Cultivation area can be more suited for an auxiliary variable than the production, because production changes every year by various factors when cultivation area doesn't change for years. The survey can be designed before harvesting by using cultivation area as an auxiliary variable.
On the other hands, another trend of material input seemed to be found by farm-size level (Table 3). Therefore, this study uses two types of estimation: ratio estimator and poststratified ratio estimator.

Table 2: Distribution and sampling of producers by farm size

	Cultivating area			*total*
	～2ha	2～5ha	Over 5ha	
Number of producers	81%	14%	5%	100%
Planting area	41%	28%	30%	100%
Sampling ratio (Number of producers)	12%	85%	95%	26%

Table 3: Coefficients of variance in material input of surveyed data

	～2ha	2～5ha	Over 5ha	*total*
Gasoline	0.97	0.94	0.68	0.95
Diesel oil	0.46	0.91	0.51	0.90
Fertilizer	0.93	0.61	0.63	1.09
Agrichemicals	0.65	0.65	0.90	1.32
N application	0.49	0.34	0.79	1.27

The ratio estimator is the amount of inputs by parent population estimated using inputs by the sample and the ratio between the auxiliary variables of the sample and the parent population. This case utilizes the cultivation area of rice as an auxiliary variable as shown as equation (1).

$$\hat{\tau}_{y_i,R} = \tau_x(\bar{y}_i/\bar{x}) \qquad (1)$$

Where, $\hat{\tau}_{y_i,R}$ is the estimated amount of input y_i, τ_x is the total cultivation area of the parent population, $\bar{y}_i$ is the average input of material i in surveyed producers, $\bar{x}$ is the average rice cultivation area in surveyed producers, and i is the type of input.

Standard error ratio of the ratio estimator is approximated as equation (2).

$$CV(\hat{\tau}_{y_i,R}) = \sqrt{\frac{\tau_x^2}{\bar{x}^2}(1-\frac{n_i}{N})\frac{1}{n_i(n_i-1)}\sum_j (y_{ij}-\hat{R}x_j)^2} \Big/ \hat{\tau}_{y_i,R} \qquad (2)$$

With $CV(\hat{\tau}_{y_i,R})$: standard error ratio of total input quantity estimation of material i, or coefficient of variance of estimates $\hat{\tau}_{y_i,R}$, n_i: number of samples in materials i, N: number of all producers, y_{ij}: total input materials i by producer j, $\hat{R}$: stands for $\bar{y}_i/\bar{x}$, x_j: rice cultivation area of producer j, j: producer..

The poststratified ratio estimator divides the sample into several strata and estimates using a ratio estimator in each stratum. In this case, stratified survey sampling has not been implemented, however, here assumes stratified sampling ex-post facto by utilizing existing sampled data. This study divided the sample into three strata by cultivation area as shown in Table 2. The equation of estimation by the poststratified estimator is shown as equation (3).

$$\hat{\tau}_{y_i,PS} = \sum_d \tau_{x,d}(\bar{y}_{i,d}/\bar{x}_d) \qquad (3)$$

Where, $\hat{\tau}_{y_i,PS}$ is the poststratified estimated amount of input y_i, $\tau_{x,d}$ is the total cultivation area of the parent population in stratum d, $\bar{y}_{i,d}$ is the average input of material i in surveyed producers of stratum d, and $\bar{x}_d$ is the average rice cultivation area in the surveyed producers of stratum d.

Standard error ratios in poststratified ratio model are also calculated.

In this case, the data representativeness of gasoline, diesel oil, fertilizer, and nitrogen fertilizer application (N_2O), and of agrochemicals, was evaluated because these data were collected by each producer surveyed. The percentages of the sample for which each input datum in the parent population was collected are presented in Table 4. Activity data related to fertilizer and agrochemicals were collected in all surveyed producers because such data on this product are managed by agricultural cooperatives to confirm cultivation standards. However, since energy consumption data have not been collected routinely, the response rate for this data was lower.

The standard error ratio of the total GHG emissions from five material inputs was estimated by Monte Carlo simulation using the standard error ratio of each material as the source of the parameters of the (normal) distribution.

In addition, optimal sampling design was considered in this case. Stratified sampling and Neyman allocation (Optimal alocation) were applied. The number of producers to survey was estimated when the confidence level was 95 %.

Table 4: Sampling ratio by input materials

	Gasoline	Diesel oil	Fertilizer	Agrichemicals	N application
Sampling ratio	16%	16%	26%	26%	26%

3.2. Results of representativeness evaluation

Table 5 indicates the standard error ratio of each material and the total GHG emissions from five materials input. The surveyed data of gasoline is considerably variable, as shown in Table 3, and its uncertainty under both estimation methods is higher than that of other input materials. However, the poststratified estimator performed better than the ratio estimator for other input materials. An especially significant effect of stratification was found in fertilizer and agrochemicals. The standard error ratio of total GHG emissions is about 3.8%, corresponding to a ± 7% confidence interval at the 95% confidence level. This survey is considered to have sufficient reliability in terms of data representativeness.

The required number of samples when implementing stratified sampling is represented in Figure 4. Tolerances are set as 1% and 5% in the 95% confidence level. This is a stricter criterion than the performance that resulted in Table 5. “Total” indicates the minimum sample size that maintains the performance set for all input materials. The number of required samples in the case of 5% tolerance is lower than in the survey actually implemented, although the accuracy is better under optimal sampling that covers smaller producers.

Table 5: Results of standard-error ratio estimations

		Gasoline	Diesel oil	Fertilizer	Agrichemicals	N_2O from N application	*Total*
Standard error ratio	ratio estimator	11.3%	7.6%	11.2%	13.8%	9.1%	5.9%
	poststratified estimator	11.3%	5.8%	5.7%	6.5%	2.3%	3.8%
Average GHG emission (kgCO_2eq/10a)	ratio estimator						131.9
	poststratified estimator						138.9

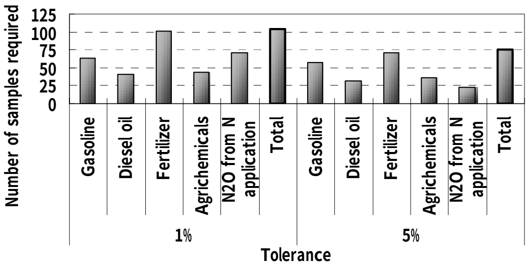

Figure 4: Number of samples required at 95% confidence level

4. Discussion

In CFP calculation, CH_4 emissions made important contributions to total GHG emissions. Although this study could not apply detailed estimation by restriction of data, it is necessary to conduct evaluation in detail including emissions models or measurements, and to make efforts to reduce emissions. The results on transportation of main products and inputs imply the potential effect of local production and consumption, and its limitations.

Although the results of analysis of data representativeness are limited to those consisting of some major input materials, this suggested the importance of implementing sample surveys on CFP for products from a large number of suppliers to improve data reliability and feasibility. This study collected data on variability of material inputs, and these data can be applicable in sample design for CFP of rice produced in situations similar to this case. Also this method can apply to the data quality evaluation in case of data deficiency among large number of producers.

Next step will be including variability of yield in evaluation for precise data quality assessment, because this study evaluated only that of input materials. Besides, further data collection to expand applicability and definition of a framework enabling simple and reliable evaluation of data representativeness will be needed.

5. References

Ajinomoto Co. Inc. (2007): Ajinomoto Group LC-CO_2 emissions factor database for food-related materials

Greenhouse Gas Inventory Office of Japan (GIO) (2009): National GHGs Inventory Report of Japan

The Japan Environmental Management Association for Industry (JEMAI) (2009a): Common background database of CFP

The Japan Environmental Management Association for Industry (JEMAI) (2009b: Reference data for PCR PA-AA-01(Rice)