Paddy and Water Environment
https://doi.org/10.1007/s10333-019-00771-w

ARTICLE

An application of multilevel model for the analysis of factors influencing paddy field productivity in the Northern Vietnamese Mountains

Huong Thi Thu Hoang[1] · Anton Van Rompaey[2] · Kim Chi Vu[3] · An Thinh Nguyen[4] · Veerle Vanacker[5]

Received: 3 December 2018 / Revised: 20 July 2019 / Accepted: 27 July 2019

Abstract
In the Northern Vietnamese Mountains, paddy fields are the main source of foods and an important indicator to define wealth in ethnic communities. This paper deals with an application of multilevel model to predict spatial patterns of paddy field productivity according to hierarchical levels such as individual fields, individual households and individual villages. The data on land use and rice productivity gathered in Sa Pa district (Lao Cai province) is used as a calibration and validation dataset. Totally, 60 households owning 136 paddy plots in four villages were interviewed. The results show that in Sa Pa, productivity of paddy fields results from human–environment interactions at different organizational levels of field, household and village. The variables collected at field and household levels play more important role in the effective use of paddy fields than the variables at village level, as they account for resp. 81.2%, 18.7% and 0.1% of the variance in productivity.

Keywords Multilevel analysis · Paddy field · Northern Vietnamese Mountains · Sa Pa district

Introduction

Agriculture worldwide faces tremendous pressure for intensification over the next 50 years to satisfy the demand of the world population increasing by one-third from 2013 to 2050 (Wu and Li 2013). Not surprisingly, within agricultural science much attention goes to a better understanding of the controlling factors of agricultural productivity, its spatial patterns and its evolution over time (Castella et al. 2005; Castella and Erout 2002; Isoda et al. 2012; Mohamed et al. 2014; Mukherjee and Kuroda 2003; Van Asselen and Verburg 2013). In most cases, it is what the determinants of crop yield and agricultural productivity are. Many studies have shown that rice productivity depends on both natural and socioeconomic factors. At macro-level, FAO and other research have described the factors affecting wetland rice production. Following are the factors contributing to the large difference in rice productivity obtained by different world's rice producing countries: geographic factors

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10333-019-00771-w) contains supplementary material, which is available to authorized users.

✉ Huong Thi Thu Hoang
huongphung114@gmail.com

Anton Van Rompaey
anton.vanrompaey@kuleuven.be

Kim Chi Vu
vukimchi@gmail.com

An Thinh Nguyen
anthinhhus@gmail.com

Veerle Vanacker
veerle.vanacker@uclouvain.be

1 Faculty of Geography, VNU-University of Sciences, 334 Nguyen Trai Street, Thanh Xuan District, Hanoi, Vietnam

2 Division of Geography, Department of Earth and Environmental Sciences, K.U. Leuven, Celestijnenlaan 200E, Bus 2409, 3001 Heverlee, Belgium

3 VNU-Institute of Vietnamese Studies and Development Sciences, Hanoi, Vietnam

4 Faculty of Development Economics, VNU-University of Economics and Business, Hanoi, Vietnam

5 Earth and Life Institute, Georges Lemaître Center for Earth and Climate Research, Université Catholique de Louvain, Place L. Pasteur 3, Bte L4.03.08, Louvain-la-Neuve, Belgium

(latitude, altitude (effect on productivity through resultant differences in temperature) (FAO 1995); climatic factors (temperature regimes, rainfall, solar radiation, day-length, winds and relative humidity) (Yoshida and Parao 1976); land and soil factors (topography, level of soil fertility) (De Datta 1981; Kassam et al. 1991); water supply factors (water deficiency, flooding with consequent crop submergence, sources of water, irrigation and drainage infrastructures) (Setter et al. 1995); farming practices (rice varieties, land preparation, methods of crop establishment, weeds, diseases, insect pests and other biological agents such as rats, birds, etc.) (Fischer 1994; Horie et al. 1994); socioeconomic factors (land tenure, credit availability and market accessibility, farmers' knowledge) (Jamin and Andriesse 1993).

At micro-level, Rasyid et al. (2016) and Alam and Effendy (2017) estimate the effect of production input on rice production and the effect of socioeconomic factors on the technical efficiency of rice farms in Indonesia. These researches based on household data and results show that seed, fertilizer, pesticide and labor significantly positively affect rice production. Socioeconomic factors such as farmer age, education and experience, the number of household members and the frequency of visiting the fields had significant positive effects on the level of technical efficiency. Research of Amer (2010) in Malaysia indicated that labor available and demand for the commodity (rice) has a strong positive relationship with the rice production volume, while land size and technology are not important factors that affect any unit of differences in the rice production volume. However, these two factors still have a positive relationship or are interpreted as contribution to the rice production level. In Uy (2016), it is estimated the impact of farmers' wages, government budget and El niño on the annual performance of rice production in the Philippines.

Obviously, rice productivity relates to the biophysical properties of the land and to the adopted agricultural practices. However, in many cases specific spatial patterns of productivity appear that is unable to link directly with individual factors. This comes from the fact that productivity results from various interaction processes that operate at different levels. At individual fields, rice productivity is controlled by slope aspect, soil quality and solar insulation. At household level, the investment capacities and the income diversification of the household affect productivity. At village level, rice productivity depends on market accessibility of the village and ethno-cultural traditions. Agricultural productivity is thus a central component of biophysical, social and economic systems acting across various scales (Van Asselen and Verburg 2013).

Within land use science, the multilevel characteristics of land use and agricultural productivity were described by international scholars (Bouwman et al. 2006; Havlík et al. 2011; McConnel and Moran 2001; Nelson 2002; Overmars and Verburg 2006; Rindfuss et al. 2004; Veldkamp et al. 2001; Walsh et al. 2001; Wise et al. 2009). The analysis of such multilevel processes and their interactions is quite challenging from statistical mathematic viewpoint. Aggregation of processes to a higher level does not straightforwardly lead to a proper representation of these higher-level processes because relations identified at the micro-level (or fine resolution) are unable to translate automatically into the same relation at the macro-level (or coarse resolution) (Easterling 1997; Jones and Duncan 1995; Robinson 1950). Disaggregating higher-level variables to the lower level also result in an exaggeration of the sample size and leads to erroneous conclusions (Overmars and Verburg 2006; Woltman et al. 2012).

Recently, Hoshino (2001) and Pan and Bilsborrow (2005) proposed a new model supporting possible tools for the analysis of multilevel phenomena in land use science. The so-called multilevel models consist of hierarchical linear regression models that enable to combine correlations within and between various spatial scale levels. They were originally developed for application in social sciences (Woltman et al. 2012) but seem to offer a high potential for land use analyses because land use data often have a very clear hierarchical structure. Moreover, multilevel modeling partly reduces the effect of spatial autocorrelation when neighboring observations are nested within one group (Overmars and Verburg 2006; Pan and Bilsborrow 2005; Polsky and Easterling 2001). Despite these promising possibility, few studies on land use incorporated multilevel models because the technique requires a multilevel land use dataset (Gray et al. 2008; Hoshino 2001; Overmars and Verburg 2006; Pan and Bilsborrow 2005; Vance and Iovanna 2006; Vanwambeke et al. 2007).

The paper aims at evaluating the potential of multilevel models to predict spatial patterns of rice productivity whereby the following levels are used: individual fields, individual households and individual villages. The data on land use and rice productivity gathered in Sa Pa district (Lao Cai province in the Northern Vietnamese Mountains) will be used as a calibration and validation dataset. We hypothesize that the productivity of paddy fields results from human–environment interactions at different organizational levels (field, household and village levels) in which decisions that are taken at household level may affect the paddy field productivity more strongly than the dynamics at village level.

Based on an overview factors affecting rice productivity and characteristics of rice cultivation in Sa Pa district, we hypothesized the factors affecting rice productivity in the study area as listed in Table 3. At field level, we hypothesized that biophysical (slope, elevation), soil characteristics of each field, frequency of visiting the fields (express through travel time from household to the field), age of paddy field (for how long the field was developed) are main drivers of rice productivity. As the study area is small and has no irrigation system (rice cultivation relies only on rainfall), so we assume that the climate factors are the same. At household level, following are hypothesized as factors contributing to the large difference in rice productivity: Residence time of household (represents the experience of household in rice cultivation), labor available, investment in paddy fields (purchase hybrid seeds, chemical fertilizers, herbicide, pesticide), ethnic identity. At village level, four variables: Distance from village center to Sa Pa town (represents accessibility to market), Tourism involvement, Forest cover (related to ability to retain water for rice fields) and Paddy fields available are considered as main drivers of rice productivity.

Study area

The data for the multilevel modeling of rice productivity are gathered in two communes of Sa Pa district: Trung Chai and Lao Chai (Fig. 1). Lao Chai commune is situated about 8 km southwest of Sa Pa town and has a total of ca. 3.927 inhabitants (General Statistics Office of Vietnam 2016). It is located inside the Hoang Lien National Park and is primarily occupied by people belonging to the Hmong ethnic group.

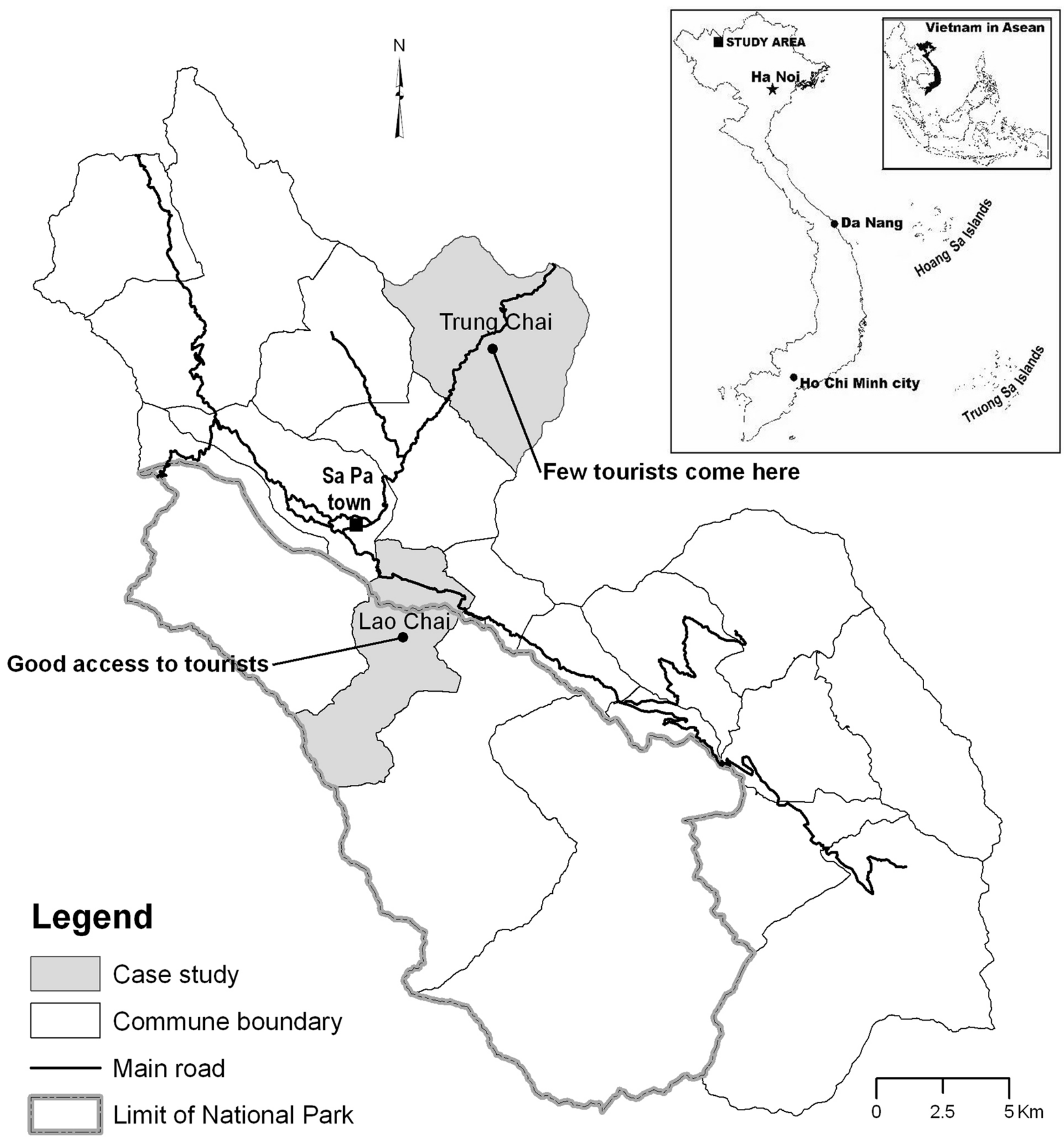

Fig. 1 Location of two selected communes in Sa Pa district, Lao Cai province of Northern Vietnamese Mountains

Lao Chai is an attractive tourism site, and ca. 22% of the local population is (partially) engaged in the tourism sector by making and selling crafts, working in restaurants, hotels or as tour guides (Hoang et al. 2014). The commune Trung Chai is located northeast of Sa Pa town (about 12 km), and has ca. 3.761 inhabitants (General Statistics Office of Vietnam 2016). It is situated outside of the National park and its population consists of people from both Hmong and Yao ethnic groups. Trung Chai is not a tourism site, and only few villagers are engaged in tourism activities (6%) (Hoang et al. 2014) (Table 1).

Paddy fields are the main source of food of Hmong and Yao livelihoods in the study area (Culas and Michaud 2004). Ethnic minorities develop terraced paddy fields on steep mountain ranges exceeding 1000 m asl. Hmong and Yao people grow subsistence rice with one crop harvest per year on the paddy fields. Due to the harsh topographic and climatic conditions, only one harvest is possible, whereas in better conditions in the lowlands two or even three harvests are possible. The rice yield is crucial for the survival of the household and an important indicator to define wealth in these ethnic communities. A household that does not produce enough rice to support its own needs is considered to be poor by the local people (Bonnin and Turner 2012).

Since 1990, the Vietnamese government has encouraged and supported modernization of agricultural practices, such as the application of chemical fertilizers and pesticides, and acquisition and sowing of high productive hybrid rice and corn breeds. This program called 'the Hunger Eradication and Poverty Reduction Programme' (HEPR, or Programme 143) aimed to increase rice productivity and improve local food security (Bonnin and Turner 2012; Turner 2012a). The agricultural modernization has led to an increase in rice production from 2.3 ton/ha in 1990 to 4.6 ton/ha in 2016 in Sa Pa district (General Statistics Office of Vietnam 2016). The increased food security resulted in an abandonment of swidden rice cultivation because of increasing land constraints, lower productivity, loss of soil fertility and lack of labor availability (Sowerwine 2004). However, the cultivation of hybrid rice requires higher financial inputs than the traditional rice cultivation. Investments are necessary for the purchase of hybrid seeds, fertilizers and pesticides. These investments are often a threshold for poor farming households to prevent them from adopting the new technologies. For households with a more diversified income from off-farm activities and cultivation of cardamom, it is relatively easier to free funds for these investments (Bonnin and Turner 2012; Turner 2007, 2012b).

Materials and methods

Data collection

Data were collected during a field trip on summer 2014 at the level of individual fields and individual households by structured interviews in four villages belonging to two communes: Lao Chai and Trung Chai. Totally 60 households were interviewed. Each household was asked to (1) answer a questionnaire on their land use practices and household characteristics and to (2) locate their paddy fields on a pan-sharpened ALOS satellite imagery with a spatial resolution of 2.5 m by 2.5 m. The satellite image was printed in color at 1:10,000 scale with annotations of major roads and public facilities (Fig. 2b).

The selection of households to be interviewed was based on systematic random sampling using the list of household heads from the Vietnam Rural, Agricultural, and Fishery Census of 2011. The questionnaire was designed as to have an exhaustive list of variables that can affect productivity of paddy fields (Online Appendix S1). The location

Table 1 Overall characteristics of the two selected communes within Sa Pa district

Name of commune	Ethnic group	Population 2016*	Accessibility	% of employment in tourism**	Located in Hoang Lien National Park?	Number of interviews	
						Interviewed households	Number of their paddy plots
Trung Chai	Hmong and Yao	3761	12 km from Sa Pa town by asphalt road	5.8	No	27	73
Lao Chai	Hmong	3927	8 km from Sa Pa town (6.5 km on asphalt road and 1.5 km on gravel road)	22.2	Yes	33	63

*National Vietnamese Census, Year book in 2016 (General Statistics Office of Vietnam 2016)

**Derived from household interviews carried out in the summer of 2014

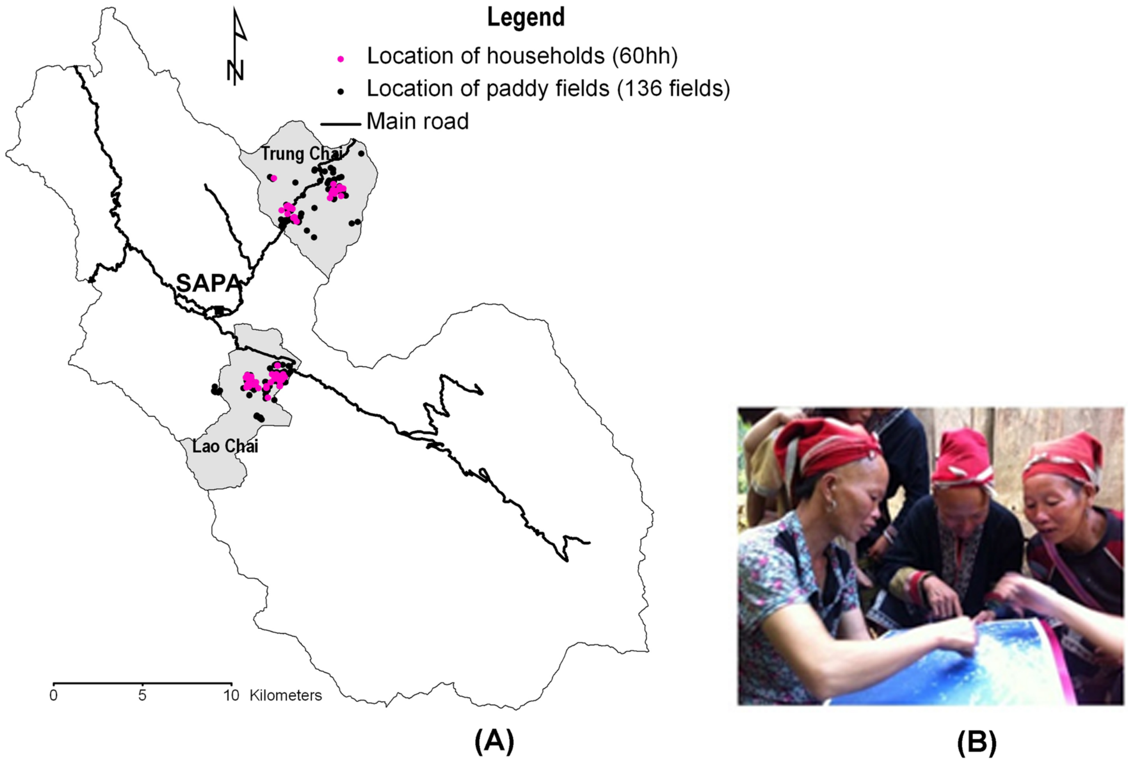

Fig. 2 **a** Location of interviewed households and their paddy fields and **b** Yao women identifying their paddy fields on the map, in Trung Chai Commune (Survey on summer 2014)

of households was mapped with handle GPS in the field. In total, 60 households and 136 paddy fields were located (Fig. 2a).

Fields were defined as a piece of land of a single owner, used for rice cultivation (Fig. 3). A household often owns a number of fields at different locations. The interviews focused on household information (occupations, composition, ethnic identity, residence time, labor available and income sources), agricultural inputs and outputs (seeds, chemical fertilizers, pesticide, herbicide, rice production, corn production and livestock), the cultivated areas, history of paddy fields and travel time from the household's residence to the fields.

Biophysical (slope, elevation) and soil characteristics of each field were determined by overlaying the location of fields on topographic maps (MONRE 2009) and soil maps (Nguyen 2007) (Fig. 4). The fields that belong to the same household will have the same household characteristics.

There are four types of soil in the study area: Gleyic Ferralsols (Fl), Humic Ferralsols formed on granite rock (HFa), Humic Alisols formed on granite rock (Ha) and Humic Alisols formed on metamorphism rock (Hj). Most paddy fields are distributed on Fl and HFa (Figs. 3 and 4). Fl soil is distributed on gentle slopes or terraced fields under 700 m above sea level. The soil profile is divided into different layers. As the top layer is flooded and plowed regularly, it has no structure with strong drainage. This layer is characterized by acidity ($pH_{KCl} = 4.8–6.0$) and low adsorption capacity (CEC < 8 me/100g). The soil layer at 40–50 cm deep still retains the characteristics of red feralit soil on acidic magma rock with small granular or lumpy structure, acidity, low absorption capacity, poor humus and total substances; the concentration of easily digestive substances is medium (Nguyen 2007). HFa is distributed at elevation from 700 to 1700 m above sea level with a slope of 8–25 degrees. It is formed on granite rock. This soil is characterized by acidity ($pH_{KCl} = 4$) and low adsorption capacity (CEC < 11.5 me/100g) (Nguyen 2007). The total and easy content of the surface substances are from poor to moderate, and they decrease gradually with depth. The natural soil quality of these two soil types are described in Table 2.

Hierarchical structure of the dataset

For the multilevel analysis, the hierarchical structure of data was constructed from the lowest level to the highest level as

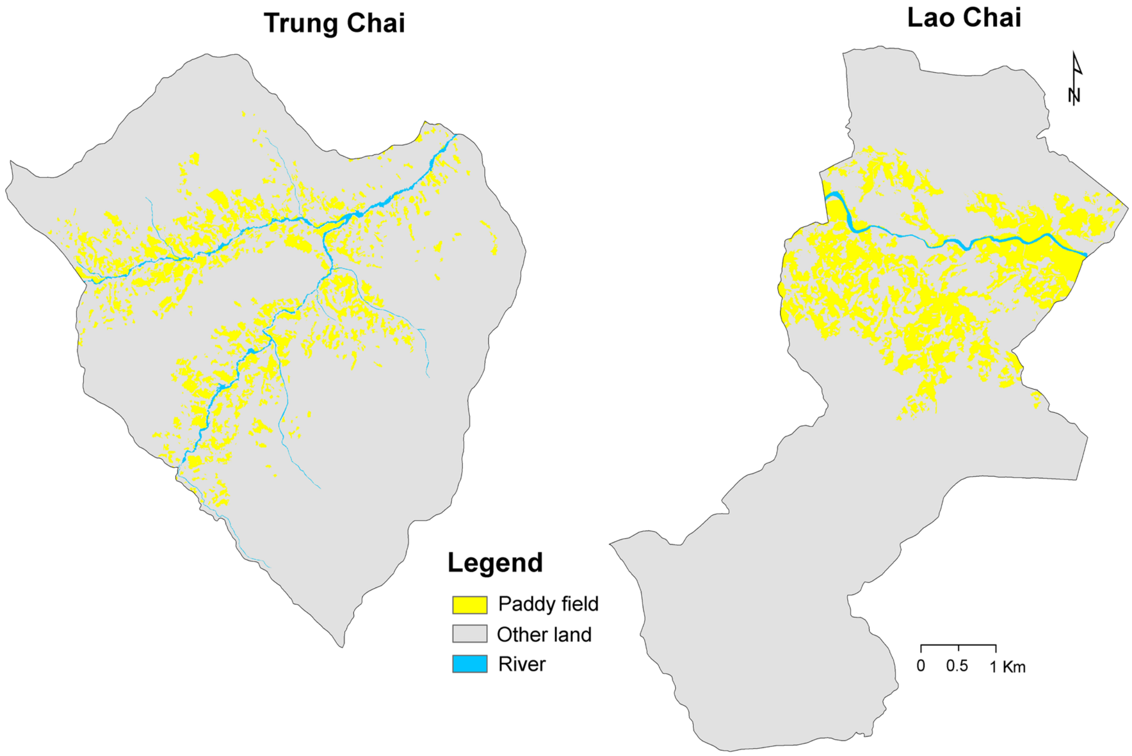

Fig. 3 Distribution of paddy fields (yellow color) in Trung Chai and Lao Chai commune in 2014 (derived from VHR-SPOT 5 image), (MONRE 2014). (Color figure online)

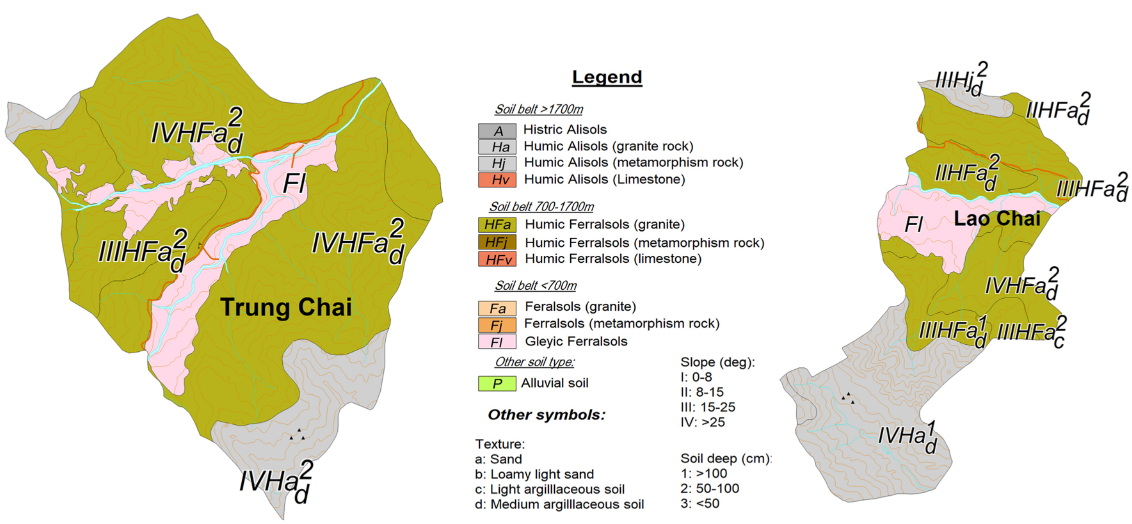

Fig. 4 Soil map of Trung Chai and Lao Chai commune (Nguyen 2007)

follows: fields—households—villages. Each variable was collected at its corresponding level. Table 3 shows the different variables that were collect at the different hierarchical levels. The rice productivity on the individual fields was considered as dependent variable.

Calibration of multilevel models for the assessment of rice productivity

A three-level model was developed and calibrated following an approach described by Overmars and Verburg (2006). The basic equation is as follows (Eq. 1):

Table 2 Results of soil quality analysis (Nguyen 2007)

Soil deep	pH_{KCl}	Humic (%)	Total (%)		Easy absorbed content (mg/100 soil gram)		
			N	P_2O_5	N	P_20_5	K_20
Fl							
1 (0–30 cm)	4.8	2.37	0.15	0.07	–	1.0	25.0
2 (30–70 cm)	6.0	1.27	0.13	0.07	3.0	1.5	12.0
HFa							
1 (0–30 cm)	4.0	3.92	0.19	0.06	4.0	1.0	23.0
2 (30–50 cm)	4.0	1.32	0.08	0.01	2.5	1.0	23.0

$$\begin{aligned} PRODUCTIVITY_{ijk} = {} & \gamma_{000} + \gamma_{100} * SLOPE\_FIELD_{ijk} + \gamma_{200} * ELEVATION\_FIELD_{ijk} \\ & + \gamma_{300} * AGE\_FIELD_{ijk} + \gamma_{400} * ACCESS\_FIELD_{ijk} + \gamma_{500} * SOIL\_FL_{ijk} \\ & + \gamma_{600} * SOIL\_HFa_{ijk} + \gamma_{010} * ETHNIC\_H_{jk} + \gamma_{020} * ETHNIC\_Y_{jk} \\ & + \gamma_{030} * RESIDENCE_{jk} + \gamma_{040} * LABOR_{jk} + \gamma_{050} * INVEST_{jk} \\ & + \gamma_{001} * TOURISM\_VILLAGE_{k} + \gamma_{002} * ACCESS\_VILLAGE_{k} \\ & + \gamma_{003} * FOREST\_VILLAGE_{k} + \gamma_{004} * PADDY\_VILLAGE_{k} + r_{0jk} + u_{00k} + e_{ijk} \end{aligned} \quad (1)$$

Whereby: $PRODUCTIVITY_{ijk}$ is the rice productivity in ton/ha for field '*i*' belonging to household '*j*' in village '*k*.' For the acronyms of the other variables, we refer to Table 3.

The coefficients for each variable are structured as follows:γ_{000} is the intercept; γ_{100}, γ_{200}, γ_{300}, γ_{400}, γ_{500}, γ_{600} are vectors of field-level coefficients; γ_{010}, γ_{020}, γ_{030}, γ_{040}, γ_{050} are vectors of household-level coefficients; γ_{001}, γ_{002}, γ_{003}, γ_{004} are vectors of village-level coefficients; r_{0jk} is the household-level random part; u_{00k} is the village-level random part; and e_{ijk} is the field-level random part.

We present four different random intercept models. The first model is the null model, which informs about the variability at the different levels. In the subsequent models, the variables will be added per level to see the influence of these groups of variables on the variance component of the higher levels. The analysis was performed with the HLM 6 software (Raudenbush et al. 2004).

To indicate the proportion of variance that is accounted for by the group level, the interclass correlation coefficients (ρ_r and ρ_u, for the household and village levels, respectively) are calculated following Eq. 2 (household level) and Eq. 3 (village level) (Browne et al. 2005; Overmars and Verburg 2006; Snijders and Bosker 1999):

$$\rho_r = \operatorname{var}(r_{0jk})/(\operatorname{var}(r_{0jk}) + \operatorname{var}(u_{00k}) + \pi^2/3) \quad (2)$$

$$\rho_u = \operatorname{var}(u_{00k})/(\operatorname{var}(r_{0jk}) + \operatorname{var}(u_{00k}) + \pi^2/3) \quad (3)$$

where ρ_r and ρ_u are the interclass coefficients for the household and village levels, var(r_{0jk}) is the variance of the random intercept at household level and var(u_{00k}) is the variance of the random intercept at village level. A logistic distribution for the level one (field level) residual implies a variance of $\pi/3$, which appears as the field-level variance in Eqs. 2 and 3 (Overmars and Verburg 2006; Snijders and Bosker 1999).

Results

Table 4 shows the calibrated values for the different stages of the multilevel model. Model 1 (Null model) is the empty model, which does not include any explanatory variables, but only includes random effects at the higher levels. Model 1 shows that the variance is significant ($p < 0.001$) at levels 2 and 3. The interclass correlation coefficients (ρ_r and ρ_u) indicate that 18.8% of the variance can be attributed to the household level and 25.6% to the village level. The remaining variance is in level 1. Thus, both the household and the village variables might be considered as predictors of productivity of paddy fields. This is studied in more detail with the models 2, 3 and 4.

Model 2 introduces a set of geographic, biophysical and soil variables at field level that are considered as explanatory variables for the productivity of paddy fields. The result shows that 'elevation,' 'soil_Fl' and 'travel time from hh to field' variables are significant. The fields that are located at higher elevations or further away from the household's residence seem to have lower productivity because of inaccessibility and less labor input. The coefficients of model 2 show that if the elevation of a field increases by 100 m, the productivity will decrease by 0.87 ton/ha. Similarly, if

Table 3 Description of the variables in the dataset used in the multilevel analysis

Variable name	Acronym	Description	Overview of factors influencing rice productivity	Unit	Min	Max	Mean	Sources
Dependent variables (field level: level 1, n = 136)								
Rice productivity	PRODUCTIVITY	Productivity of each paddy field		Ton/ha	1	10	4	Interview in 2014
Independent variables at field level (level 1, n = 136)								
Slope	SLOPE_FIELD	Slope of paddy field that was determined by overlaying the location of field on topographic map	Topography interferes with the distribution of water supplies (FAO 1995). Rice paddies on steeper slopes and higher altitudes are generally less productive (Isoda et al. 2012)	Degree	0	35	17	Topographic map (MONRE 2009)
Elevation	ELEVATION_FIELD	Elevation of paddy field that was determined by overlaying the location of field on topographic map		m. asl	800	1345	1099	
Travel time from household to the field	ACCESS_FIELD	Travel time from the residence of the household to the field	Travel time from household to the field affects frequency of visiting the fields. Higher frequency visiting the fields led to increased levels of technical efficiency for rice farms (Rasyid et al. 2016)	Minute	1	150	15	Interview in 2014
Age of field	AGE_FIELD	For how long the field was developed	Newer fields maybe not less productive than older ones (Isoda et al. 2012)	Year	2	103	54	
Soil_Fl	SOIL_FL	1 if soil type is Fl, 0 otherwise	Yield of wetland rice is influenced by the level of soil fertility (Alam and Effendy 2017). The deficiency of nitrogen, phosphorus, humus are common constraint to wetland rice production (Cassman et al. 2005; FAO 1995; Jamin and Andriesse 1993)	0 or 1	0	1	0.65	Extraction from Soil map of Sa Pa district (Nguyen 2007)
Soil_HFa	SOIL_HFa	1 if soil type is Hfa, 0 otherwise		0 or 1	0	1	0.33	

Table 3 (continued)

Variable name	Acronym	Description	Overview of factors influencing rice productivity	Unit	Min	Max	Mean	Sources
Independent variables at household level (level 2, n = 60)								
Residence time	RESIDENCE	Number of generations for that the family have been living in the Sa Pa	Age of farmers or residence time was positively correlated with rice farm experience and increased experience and knowledge led to increased technical efficiency (Rasyid et al. 2016)	Generation	2	7	4	Interview in 2014
Labor	LABOR	Number of household individuals from 15 to 60 years old	Additional labor allows rice farm activities, for example, soil tillage, weeding, fertilization, pest and disease control, to be implemented in a timely manner that in turn will increase rice production (Li et al. 2008; Alam and Effendy 2017; Rasyid et al. 2016)	Person	2	7	4	
Investment in paddy fields	INVEST	Total investment of household in paddy fields (purchase hybrid seeds, chemical fertilizers, herbicide, pesticide, etc.)	Fertilizer had a significant effect on rice production. Additional fertilizer on agricultural land increased the nutrients nitrogen, sulfur and potassium in the soil that were needed by rice (Li et al. 2008; Alam and Effendy 2017; Rasyid et al. 2016). Pesticides significantly affected rice production, reducing damage to rice panicles by pests and disease, to maintain production (Dewi and Idris, 2005)	kVND/ha/year	3202	10,931	5691	
Ethnicity Hmong	ETHNIC_H	1 if household head is Hmong, 0 otherwise	Hmong have their paddy fields on steeper slopes at higher altitudes compared to others ethnicity (Jadin et al. 2013) so maybe they get less rice productivity	0 or 1	0	1	0.9	

Table 3 (continued)

Variable name	Acronym	Description	Overview of factors influencing rice productivity	Unit	Min	Max	Mean	Sources
Ethnicity Yao	ETHNIC_Y	1 if household head is Yao, 0 otherwise		0 or 1	0	1	0.1	
Independent variables at village level (level 3, n = 4)								
Distance to district capital	ACCESS_VILLAGE	Distance from village center to Sa Pa town	Distance from village center to Sa Pa town represents accessibility to market for buying rice input	km	8	12	10	Topographic map (MONRE 2009)
Tourism involvement	TOURISM_VILLAGE	% of the population of the village involved in tourism industry	The introduction of tourism increased the living standard of the ethnic minorities and led to more intensive farming systems with forest regrowth on abandoned fields (Hoang et al. 2018)	%	1	25	8	Interview in 2014
Forest cover	FOREST_VILLAGE	Percentage of area covered by forest	The agro-environment affected by deforestation had an impact on cropland stability and productivity (Lim et al. 2017).	%	19	53	40	Land cover map 2014
Paddy field available	PADDY_VILLAGE	Average paddy field area used per household	Land area/household can explained variation in the application of rice technology by farmers (FAO 1995)	ha/household	8	15	12	

Table 4 Multilevel models for productivity of paddy fields

Paddy productivity (ton/ha)	Model 1 (Null model)	Model 2 (level 1)	Model 3 (level-2)	Model 4 (level-3)
Fixed effects				
Level 1 (field level)				
Intercept	4.166392**	12.766173**	6.679641*	7.339103*
Slope		0.055026	0.050709	0.054379
Elevation		−0.008680***	−0.005380*	−0.005777*
Travel time from household to field		−0.016496*	−0.014409*	−0.014677*
Age of field		−0.001633	0.001400	0.001253
Soil_Fl		0.204652	0.150494	0.368720
Soil_HFa		−0.204652*	−0.150494*	−0.368720*
Level 2 (household level)				
Investment in agriculture			0.000242*	0.000241*
Labor			0.016421	0.008247
Residence time			0.257579*	0.259640*
Ethnic Hmong			0.073783	0.151141
Ethnic Yao			−0.093229	−0.171557
Level 3 (village level)				
Distance to district capital				−0.000087
Tourism involvement				0.06733
Forest cover				0.043470*
Paddy field available				0.050572
Random effects				
Level 2				
var(r_{0jk})	1.10811***	0.97961***	0.57996***	0.75443***
ρ_r	0.188	0.230	0.150	0.187
Level 3				
var(u_{00k})	1.51158***	0.00038*	0.00026*	0.00028*
ρ_u	0.256	0.000089	0.000067	0.00063

*$p < 0.1$

**$p < 0.01$

***$p < 0.001$

the travel time from the household's residence to the field increases by 10 min, the productivity will decrease by 0.16 ton/ha. The fields located on HFa soil seem to have lower productivity compared to that on Fl soil. The levels 2 and 3 random parts are still significant, but the variance component of level 2 increases while the variance component of level 3 decreases substantially compared to the null model. This means that the variables included account for the household-level variability.

Model 3 adds household variables to model 2. The result shows that the productivity of paddy fields is positively related with the investment capacity in paddy fields and residence time. The coefficients of model 3 show that if the investment in paddy field increases by 1000 kVND/ha, the productivity will increase by 0.24 ton/ha. Similarly, if the residence time of household in Sa Pa district increases by 1 generation, the productivity will increase by 2.6 ton/ha. After including the household-level variables, the random part of level 2 (household level) decreased slightly. Apparently, the variance at level 2 is largely captured by the included variables. The level 1 variables are still significant. The level 3 variance decreased in comparison with model 2.

Model 4 adds the village-level variables to the model 3. The result shows that the productivity of paddy fields is positively related with forest cover. If forest cover increases by 10%, the productivity will increase by 0.43 ton/ha. Including village variables results in a higher random part at household level, while the random part at village level decreased. The interclass correlation coefficients of full model (ρ_r and ρ_u in Table 4, model 4) indicate that variables at household level contributed for 18.7% of the variance of outcome variable (rice productivity) and village variables contributed for only 0.1%. The remaining variance (81.2%) is in level 1.

Discussion

At field level, the results show that fields that are located at higher elevations or further away from the household's residence seem to have lower productivity. This can be explained by inaccessibility and less labor input. On the contrary, if the paddy fields are close to the household's residence or accessible, farmers will visit the field more often, thus leading to increased levels of technical efficiency for rice farms (Rasyid et al. 2016). Our results show that slope variable does not influence rice productivity significantly. This is consistent with research of Isoda et al. (2012) in Sa Pa district reporting that rice paddies on steeper slopes and higher altitudes are generally less productive, but that some villages can obtain higher productivity because of effective use of fertilizers. In addition, the paddy fields in our study area were created into the terraces on slope hills to reduce the effect of slope. This cultivated technique was first introduced by the Hmong and Yao who migrated from southern China to northern Vietnam in the late nineteenth and early twentieth centuries (Michaud 1997). The results show that there is no linkage between the age of the fields and productivity. This is consistent with research of Isoda et al. (2012) who reported that newer fields are not less productive than older ones, as the productivity also depends on effective use of fertilizers. The research indicates that the fields located on HFa soil seem to have lower productivity compared to the fields on Fl soil. This can be explained by the fact that HFa is distributed on steeper slopes at higher altitudes that are not very suitable for rice cultivation.

At household level, the results support the conclusions of earlier research of Hoang et al. (2018) in Sa Pa, as it shows that the productivity of paddy fields is positively related with the investment capacity of households. This research showed that households that are engaged in tourism get higher productivity because they invest more in chemical fertilizers and hybrid seeds. This was not observed here at village level, but can be due to the limited number of observations at village level ($n=4$), that is simply not enough to draw statistically sound conclusions. Our results also agree with research of Rasyid et al. (2016), Alam and Effendy (2017), Effendy (2015) and Li et al. (2008) that indicated that fertilizer had a significant effect on rice production. Additional fertilizer on agricultural land increased the nutrients nitrogen, sulfur and potassium in the soil that were needed by rice. Pesticides significantly affected rice production, reducing damage to rice panicles by pests and disease, to maintain production (Dewi and Idris 2005). Our results show that the family that has been living in the Sa Pa for longer time can harvest higher productivity. The residence time reflects farm lifecycles in indigenous land use. The households that have been living in the study area for many generations will get higher productivity as they occupy more fertile fields and have more experience with paddy field cultivation. The results agree with research of Rasyid et al. (2016) that indicates that if a farmer consistently grows rice, they will increase technical efficiency over time.

At village level, the productivity of paddy fields is positively related with forest cover. That means that the more forests, the higher rice productivity or in other words, the less forest, the lower productivity. This result supports research of Lim et al. (2017) that identify the changes in agricultural quality under deforestation such as wind erosion, water erosion, organic carbon loss, and runoff. The agro-environment affected by deforestation had an impact on cropland stability and productivity

This study suggests a bottom-up approach for improving paddy field productivity instead of a top-down approach. Investments should focus on improving the field conditions and the farming capacity of the household rather than investing in the macro-level.

Conclusion

Multilevel analysis is an interesting approach to deal with scale and level issues in land use studies. This case study has shown that the productivity of paddy fields can be explained by variables at three organizational levels: field, household and village. At the field level, 'elevation,' 'soil_HFa' and 'travel time from household's residence to field' are significant predictors of productivity. The fields that are located at higher elevations or further away from the household's residence seem to have lower productivity because of inaccessibility and less labor input. The fields on HFa soil can get lower productivity than that on Fl soil. At household level, the results show that the productivity of paddy fields is positively related with the investment capacity in paddy fields and the residence time. The households that invested more capital in paddy fields by buying hybrid seeds, chemical fertilizers and pesticides generally get higher productivity. Also, households that have been living in the study area for many generations use to get higher productivity as they occupy the more fertile fields and have more experience in paddy field cultivation. At village level, the productivity of paddy fields is positively related to forest cover.

The results of multilevel modeling reveal at which organizational levels the explanatory variables explain the land use intensification. Productivity of paddy fields is predominantly determined by geographic, biophysical and variables at field level, that account for 81.2% of the variance in productivity. Variables at village level explain only 0.1% of the outcome variable. The variables that were collected at household level also play an important role and account for 18.7% of the variance in productivity. These results suggest that the decisions that are taken at household level may affect the land use more strongly than the dynamics at village level.

Acknowledgements This research is funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED). We are thankful to prof. Yuzuru ISODA, prof. Doo-chul KIM for sharing their expertise on household surveys in Northern Vietnam. We also thank village leaders and local people in Sa Pa district for facilitating the interviews and the collection of field data.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Alam MN, Effendy (2017) Identifying factors influencing production and rice farming income with approach of path analysis. Am J Agric Biol Sci 12:39–43

Amer A (2010) A study on factors affecting rice production in malaysia. Paper presented at the 1st. international conference on arts, social sciences and technology iCAST 2010, Pulau Pinang, Malaysia

Bonnin C, Turner S (2012) At what price rice? Food security, livelihood vulnerability, and state interventions in upland northern Vietnam. Geoforum 43:95–105

Bouwman AF, Kram T, Klein Goldewijk K (2006) Integrated modelling of global environmental change. An overview of IMAGE 2.4. Netherlands Environmental Assessment Agency

Browne WJ, Subramanian SV, Jones K, Goldtein H (2005) Variance partitioning in multilevel logistic models that exhibit over-dispersion. J R Stat Soc A 168:599–613

Cassman K et al (2005) Cultivated systems. In: Scholes R, Rashid H (eds) Millennium ecosystem assessment: working group on conditions and trends. Island Press, Washington, pp 745–794

Castella JC, Erout A (2002) Montaine paddy rice: the cornerstone of agricultural production systems in Bac Kan province Vietnam. In: Castella JC, Quang DD (eds) Doi Moi in the Mountains. Land use changes and farmers's livelihood stratergies in Bac Kan province Vietnam. The Agricultural Publishing House, Hanoi, pp 175–195

Castella J-C, Boissaua S, Trung TN, Quang DD (2005) Agrarian transition and lowland-upland interactions in mountain areas in northern Vietnam: application of a multi-agent simulation model. Agric Syst 86:312–332

Culas C, Michaud J (2004) A contribution to the study of the Hmong (Miao) migrations and history. In: Tapp N, Michaud J, Culas C, Lee GY (eds) Hmong-Miao in Asia. Silkworm Books, Chiang Mai, pp 71–96

De Datta SK (1981) Principles and practices of rice production. Wiley, Amsterdam

Dewi S, Idris (2005) Production efficiency of rice farms system in rice field of technical irrigation. Asian J Agric Ext Econ Sociol 25:1–12

Easterling WE (1997) Why regional studies are needed in the development of full-scale integrated assessment modelling of global change process. Glob Environ Change Part A 7:337–356

Effendy (2015) Increasing of cocoa farmers household income with two stage least squares method. Mod Appl Sci 9:120–127. https://doi.org/10.5539/mas.v9n6p120

FAO (1995) Factors affecting wetland rice production and the classification of wetlands for agricultural production. http://www.fao.org/3/x6611e/x6611e03c.htm. Accessed 10 May 2019

Fischer KS (1994) New breakthroughs and present accomplishments in rice research in Asia. Paper presented at the 18th session of the international rice commission, Rome, Italy, 4–9 September 1994

General Statistics Office Of Vietnam (2016) National Vietnamese census year book. Statistical Publishing House, Lao Cai

Gray CL, Bilsborrow RE, Bremner JL, Lu F (2008) Indigenous land use in the ecuadorian amazon: a cross-cultural and multilevel analysis. Hum Ecol 36:97–109

Havlík P et al (2011) Global land-use implications of first and second generation biofuel targets. Energy Policy 39:5690–5702

Hoang HTT, Vanacker V, Van Rompaey A, Vu KC, Nguyen AT (2014) Changing human-landscape interactions after development of tourism in the northern Vietnamese Highlands. Anthropocene 5:42–51

Hoang TTH et al (2018) Impact of tourism development on the local livelihoods and land cover change in the Northern Vietnamese Highlands. Environ Dev Sustain. https://doi.org/10.1007/s10668-018-0253-5

Horie T, Ohrishi M, Angus JF, Lewin LG, Matano T (1994) Physiological characteristics of high yielding rice inferred from cross-location experiment. Paper presented at the temperate rice achievements and potential, Yanco, Australia, 21–24 Feb 1994

Hoshino S (2001) Multilevel modeling on farmland distribution in Japan. Land Use Policy 18:75–90

Isoda Y, Nguyen NH, Kanda T, Yotsumoto Y, Nguyen AT, Hoang TTH, Kim D-C (2012) Terraced paddy expansion by ethnic minorities in Sapa, Vietnam: identifying locations and measuring productivity. Paper presented at the subtainability of rural systems in global economic and environmental issues, IGC, University of Cologne

Jadin I, Vanacker V, Hoang TTH (2013) Drivers of forest cover dynamics in smallholder farming systems: the case of Northwestern Vietnam. Ambio 42:344–356

Jamin JY, Andriesse W (1993) Discussion synthesis. Paper presented at the annual workshop of Inland Valley Consortium, WARDA, Bouake, Côte d'Ivoire, 8–10 October

Jones K, Duncan C (1995) Individuals and their ecologies: analysing the geography of chronic illness within a multilevel modelling framework. Health Place 1:27–40

Kassam AH, Van Velthuizen HT, Fischer GW, Shah MM (1991) Agro-ecological land resources assessment for agricultural development planning: a case study of Kenya. FAO and International Institute for Applied Systems Analysis

Li X, Luo Y, Gao Q, Dong S, Yang X (2008) Farm production growth in the upper and middle parts of the yellow river basin, China, during 1980–1999. Agric Sci 7:344–355

Lim C-H, Choi Y, Kim M, Jeon SW, Lee W-K (2017) Impact of deforestation on agro-environmental variables in cropland, North Korea. Sustainability 9:1354

McConnel W, Moran EF (2001) Meeting in the middle: the challenge of meso-level integration. Paper presented at the an international workshop on the harmonization of land use and land cover classification. LUCC Report Series No. 5, Indiana University, Bloomington, USA

Michaud J (1997) From Southwest China into upper Indochina: an overview of Hmong (Miao) migration. Asia Pacific Viewpoint 38:119–130

Mohamed ES, Saleh AM, Belal AA (2014) Sustainability indicators for agricultural land use based on GIS spatial modeling in North of Sinai-Egypt. Egypt J Remote Sens Space Sci 17:1–15

MONRE (2009) Topographic maps at the scale 1:50,000. Ministry of Natural Resources and Environment

MONRE (2014) VHR-SPOT image 2014. Remote Sensing Center, Ministry of Natural Resources and Environment, Vietnam

Mukherjee AN, Kuroda Y (2003) Productivity growth in Indian agriculture: Is there evidence of convergence across states? Agric Econ 29:43–53

Nelson GC (2002) Introduction to the special issue on spatial analysis for agricultural economists. Agric Econ 27:197–200

Nguyen AT (2007) Analyzing structure of landscape ecology for sustainable development of agriculture, forestry and tourism in Sa Pa district, Lao Cai province. VNU University of Science, Hanoi

Overmars KP, Verburg PH (2006) Multilevel modelling of land use from field to village level in the Philippines. Agric Syst 89:435–456

Pan W, Bilsborrow RE (2005) The use of a multilevel statistical model to analyze factors influencing land use: a study of the Ecuadorian Amazon. Glob Planet Change 47:232–252

Polsky C, Easterling WE (2001) Adaptation to climate variability and change in the US great plains: a multi-scale analysis of Ricardian climate sensitivities. Agric Ecosyst Environ 85:133–144

Rasyid MN, Setiawan B, Mustadjab MM, Hanani N (2016) Factors that influence rice production and technical efficiency in the context of an integrated crop management field school program. Am J Appl Sci 13:1201–1204

Raudenbush S, Bryk A, Congdon K (2004) HLM 6, Hierarchical linear and non-linear modelling. SSI (Scientific Software International), Skokie

Rindfuss RR, Walsh SJ, Turner BL II, Fox J, Mishra V (2004) Developing a science of land change: challenges and methodological issues. PNAS 101:13976–13981

Robinson W (1950) Ecological correlations and the behaviour of individuals. Am Sociol Rev 15:351–357

Setter TL, Ingram KT, Tuong TP (1995) Environmental characterization requirement for strategic research in rice grown under adverse conditions of drought, flooding, or salinity. In: Ingram IKT (ed) Rainfed lowland rice: agricultural research for high-risk environment. IRRI, Manila, pp 3–18

Snijders TAB, Bosker RJ (1999) Multilevel analysis: an introduction to basic and advanced multilevel modelling. Sage, New York

Sowerwine J (2004) The political ecology of Dao (Yao) landscape transformations: territory, gender, and livelihood politics in Highland Vietnam. University of California, Berkeley

Turner S (2007) Trading old textiles: the selective diversification of Highland livelihoods in northern Vietnam. Hum Organ 66:401–416

Turner S (2012a) "Forever hmong": ethnic minority livelihoods and Agrarian transition in upland Northern Vietnam. Prof Geogr 64:540–553

Turner S (2012b) Making a living the hmong way: an actor-oriented livelihoods approach to everyday politics and resistance in upland Vietnam. Ann Assoc Am Geogr 102:403–422

Uy PJ (2016) Economic factors affecting rice production in the Philippines. College of Arts and Sciences

Van Asselen S, Verburg P (2013) Land cover change or land-use intensification: simulating land system change with a global-scale land change model. Glob Change Biol 19:3648–3667. https://doi.org/10.1111/gcb.12331

Vance C, Iovanna R (2006) Analyzing spatial herarchies in remotely sensed data: insights from a multilevel model of tropical deforestation. Land Use Policy 23:226–236

Vanwambeke S, Somboon P, Lambin E (2007) Rural transformation and land use change in Northern Thailand. J Land Use Sci 2:1–29

Veldkamp A, Kok K, De Koning GHJ, Schoorl JM, Sonneveld MPW, Verburg PH (2001) Multi-scale system approaches in agronomic research at the landscape level. Soil Tillage Res 58:129–140

Walsh SJ, Crawford TW, Welsh WF, Crews-Meyer KA (2001) A multi-scale analysis of LULC and NDVI variation in Nan Rong district, northeast Thailand. Agric Ecosyst Environ 85:47–64

Wise M et al (2009) Implications of limiting CO_2 concentrations for land use and energy. Science 324:1183–1186

Woltman H, Feldstain A, MacKay JC, Rocchi M (2012) An introduction to hierarchical linear modeling. Tutor Quant Methods Psychol 8:52–69

Wu J, Li M (2013) Land use change and agricultural intensification: key research questions and innovative modeling approaches. The International Food Policy Research Institute, Washington

Yoshida S, Parao FT (1976) Climatic influence on yield and yield components of lowland rice in the tropics. Int Rice Res InstClimate Rice. https://eurekamag.com/research/000/615/000615716.php