Supplementary Material S3: Model parameters for land use system change

Abstract

Summary of the model parameters used in CLUMondo to derive plausible land use systems for 2030. The CLUMondo model consists of two separate components: a demand module that does not consider spatial factors, and an allocation module that accounts for spatial considerations. The demand module evaluates alterations in the demand for ecosystem goods and services at a global level, which are then transformed into changes in land usage in particular areas through the allocation module. The scenarios were based on sets of demands for commodities and services from existing policy decisions, which reflect the main land use orientations of the development plans for the Mekong Delta: scenario "Specialisation" (SPE) for the pair of decisions from 2012 No.124/QD-TTg and No. 939/QD-TTg; and scenario "Diversification" (DIV) for the pair of decisions from 2014 and 2018 No. 639/QD-BNN-KH and No. 816/QD-BNN-KH.

Contents

Local suitability of land use systems

At any given time, land use system changes are anticipated to occur in areas that are best suited for particular land use systemstypes. Suitability, in this context, is determined by the interplay between relevant actors and local conditions. To gauge the suitability of locations, various biophysical and socioeconomic factors that are deemed to influence land use system changes were evaluated empirically.

1.1 Suitability layers

In order to identify and measure the associations between particular land use system types and a set of explanatory factors, also known as covariates, a statistical analysis is employed. The selection of covariates was determined with key informants, stakeholders and local communities by discussing the factors that contribute to land use change in the area. For instance, elevation, proximity to roads, and the salinity content of soil may be regarded as significant determinants of cropland distribution in the Mekong Delta, and hence are the covariates that would be incorporated into the statistical analysis.

Table 1 Suitability layers used to predict likelihood of land system spatial occurrence

Topic	Layer	Type	Year	Reference		Justification
Geomorphology	Elevation (DEM)	Static	2013	Robinson et al. (2014)	EarthEnv	Influences land system suitability and crop patterns
Climate	Temperature	Dynamic	2020-2030	WorldClim (2020)	WorldClim	Climate variables having a strong conditioning on the feasability
	Solar radiation	Static	2021	GSA 2.6 (2021)	Global Solar Atlas	of conducting crops
	Precipitation	Dynamic	2020-2030	WorldClim (2020)	WorldClim	
Soil	Organic carbon stock	Static	2019	ISRIC (2019)	SoilGrids	Influences land system suitability and crop patterns. Cash crops
	Clay content	Static	2019	ISRIC (2019)	SoilGrids	and intensive rice cultivation need rich soils, with a balanced clay-
	Sand content	Static	2019	ISRIC (2019)	SoilGrids	sand content. Sandy soil is associated with better drainage, which
	Cation Exchange Capacity	Static	2019	ISRIC (2019)	SoilGrids	might be positive for cash crop and fruit tree production.
	Soil organic carbon content	Static	2019	ISRIC (2019)	SoilGrids	
	Soil pH	Static	2019	ISRIC (2019)	SoilGrids	
	Soil class	Static	2019	ISRIC (2019)	SoilGrids	
	Soil depth	Static	2019	ISRIC (2019)	SoilGrids	
	Salinity	Dynamic	2020-2030	Eslami et al. (2019)	*Scientific Reports*	Positively associated with aquaculture, negatively with fruit trees
Hydrology	Water logging	Static	2016	DLR (2017)	WISDOM	Associated with aquaculture and rice cultivation
	Flood compartments	Static	2020	SIWRR (2020)	SIWRR	Dyke system controls maritime entries and flood management
Socio-economics	Population density	Static	2020	WorldPop (2020)	*Scientific Data*	Associated with urban areas
	Distance to major roads	Static	2019	Nelson et al. (2019)	*Scientific Data*	Proximity associated with urban or high value added agriculture
	Distance to major road intersections	Static	2019	Nelson et al. (2019)	*Scientific Data*	Proximity associated with urban or high value added agriculture
	Distance to major waterways	Static	2019	Nelson et al. (2019)	*Scientific Data*	Proximity associated with export agriculture/aquaculture
	Market accessibility	Static	2019	Nelson et al. (2019)	*Scientific Data*	Associated with export agriculture/aquaculture
	Market influence	Static	2019	Nelson et al. (2019)	*Scientific Data*	Associated with export agriculture/aquaculture

1.2 Regression parameters

Prior to conducting the regression analysis, we examined the covariates for correlation. When the correlation between two covariates was deemed too strong ($> .7$), one of the covariates was excluded from the analysis. The coefficients of the logit model were estimated by means of a logistic regression analysis, with the initial land use map serving as the dependent variable. Any variables that did not make a significant contribution to explaining the land use pattern were excluded from the final regression equation. To run the regression, we used a balanced sample with a size of 30% of all observations and a minimum distance between samples of 2 cells.

Logistic regression analysis cannot be evaluated using the R^2 measurement for the goodness of fit, which is commonly used in ordinary regressions. Instead, the AUC method can be used to evaluate the goodness of fit by comparing predicted probabilities with observed

Table 2 Regression coefficients for the proability model of local suitability of the different land use systems

Parameters	TripleRice	DoubleRice	CashCrops	Fruits	AquaForest	Aquaculture	RiceAqua	Urban
Constant	19.46	-36.26	36.20	127.09	-27.51	-29.24	-156.60	-10.19
Elevation	-0.23	n/a	n/a	0.10	n/a	n/a	-0.25	n/a
Temperature	-0.03	n/a	-0.06	-0.06	0.02	0.04	n/a	n/a
Solar radiation	n/a	0.78	n/a	-3.23	n/a	n/a	6.20	n/a
Precipitation	-3.97	3.60	-9.41	-8.29	5.41	5.13	-5.09	n/a
Soil organic carbon stock	n/a	n/a	n/a	-0.08	-0.04	0.03	n/a	n/a
Clay content	-2.68	-0.83	2.00	n/a	1.49	2.99	2.46	n/a
Sand content	n/a	1.54	n/a	n/a	-2.86	n/a	0.77	n/a
Cation exchange capacity	n/a	n/a	3.82	n/a	n/a	n/a	n/a	n/a
Soil organic carbon content	n/a	0.43	n/a	n/a	n/a	0.73	n/a	n/a
Soil pH	0.36	1.28	3.08	0.34	n/a	n/a	-1.41	n/a
Soil class	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Soil depth	n/a	1.28	n/a	n/a	n/a	0.00	0.00	n/a
Salinity	-0.05	-0.14	n/a	n/a	0.03	n/a	n/a	n/a
Water logging	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Flood compartments	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Population density	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.20
Market accessibility	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Market influence	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
AUC	0.817	0.810	0.840	0.864	0.802	0.840	0.899	0.888

values across the entire range of predicted probabilities. The AUC method calculates the area under the curve of a plot that shows the sensitivity of the logistic regression function versus its specificity. The AUC value ranges from .5 to 1, with .5 representing a completely random model and 1 representing a perfect fit. While there are no specific cut-off values for determining the acceptability of a logit model, values above .9 are considered excellent, values between .8 and .9 are good, values between .7 and .8 are fair, values between .6 and .7 are poor, and values below .6 are not good.

1.3 Probability maps

By taking into account the biophysical and socioeconomic factors present in a particular area, it was possible to calculate the relative likelihood of encountering a specific land use type in that location through the use of a binomial logit model. A binomial logit model expresses the probability of encountering a particular land use type at a given location in comparison to not encountering that same land use type at that location. Thus, a separate logit model was generated for each land use type.

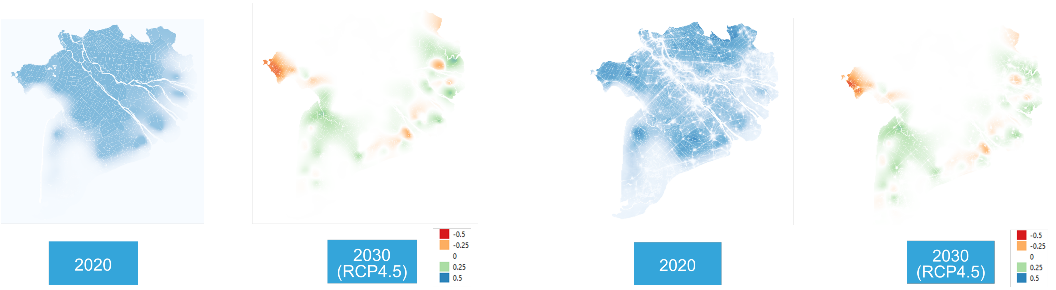

Figure 1 Probability map of triple rice (left) and double rice (right) occurrence in 2020 and projected probability change in 2030 based on the scenario RCP4.5

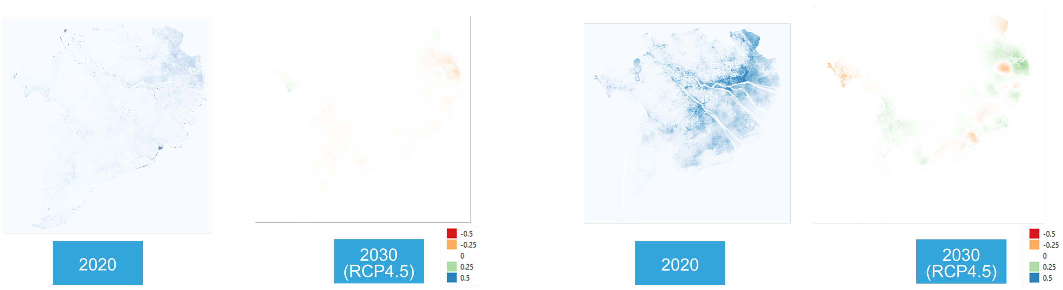

Figure 2 Probability map of cash crops (left) and fruit trees (right) occurrence in 2020 and projected probability change in 2030 based on the scenario RCP4.5

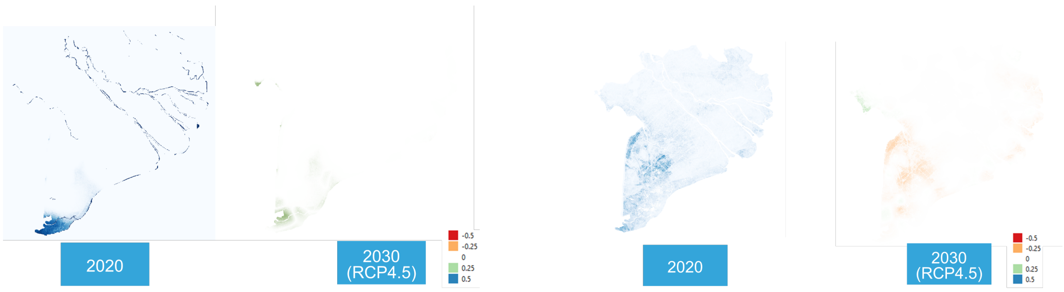

Figure 3 Probability map of the occurrence of the systems aquaculture-forest (left) and rice-aquaculture (right) in 2020 and projected probability change in 2030 based on the scenario RCP4.5

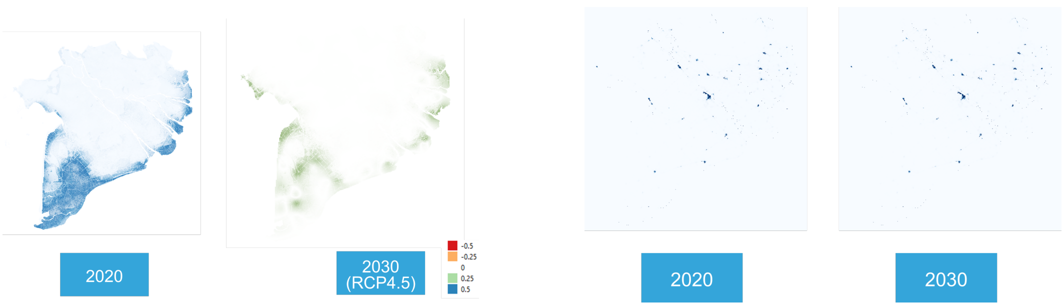

Figure 4 Probability map of aquaculture (left) and urban (right) occurrence in 2020 and projected probability change in 2030 based on the scenario RCP4.5

Land use system services

CLUMondo is primarily driven by demands for ecosystem goods and services, which ultimately determine land use changes. For instance, a demand for shrimps may result in an expansion of aquaculture areas, while a demand for housing may increase the urban footprint of a region. Sometimes, demands are directly linked to a specific land use, whereas in other cases, different land use types can provide the same ecosystem goods or services. For example, both triple rice agriculture and double rice agriculture can produce rice, albeit at varying levels of intensity. These demands were derived from trend extrapolation based on national statistics.

2.1 Land use system services matrix

One land use system can generate various ecosystem goods or services, and one particular service can be produced by different land use systems. For instance, rice-aquaculture and triple rice can both produce crops. Moreover, because there is no longer flooding, triple rice land use system can include housing units, which can contribute to the demand for crop production and residential areas. The land use systems matrix stores information about the connection between specific land uses and their production. Each row in the matrix displays the area, goods, or services provided by a specific land use system. Demand types and land use systems are related one-to-one in their simplest form, with demands given as areas for a specific land use system.

Table 3 Services associated with land use systems under scenario "Specialisation" (left) and "Diversification" (right)

SPE	Built-up	RiceProd	AquaProd
TripleRice	0.06	3.00	0.02
DoubleRice	0.04	2.00	0.12
CashCrops	0.07	0.80	0.06
Fruits	0.07	0.60	0.06
AquaForest	0.02	0.02	0.80
Aquaculture	0.04	0.02	2.00
RiceAqua	0.03	1.00	1.20
Urban	1.20	0.00	0.00

DIV	Built-up	AgriProd	AquaProd
TripleRice	0.06	1.30	0.02
DoubleRice	0.04	1.00	0.12
CashCrops	0.07	2.20	0.06
Fruits	0.07	1.70	0.06
AquaForest	0.02	0.02	1.40
Aquaculture	0.04	0.02	2.20
RiceAqua	0.03	0.60	1.80
Urban	1.20	0.00	0.00

2.2 Future demand for services

The conversion of land use system is often caused by the need for ecosystem goods and services. To illustrate, an upsurge in the requirement for food leads to an expansion of farmland in a region. Demand is described as a region designated for a particular land use system, such as a specific number of hectares for cropland, grassland, and urban land. However, land use system can also be defined by different land use intensities, and the intensity of a land use can determine the amount of land needed to meet a demand. For example, a given need for crops can be fulfilled by a small area and high intensity or by a larger area with a relatively low intensity. Furthermore, a single land-use type can generate multiple ecosystem goods and services. For instance, villages can offer both residential space and agricultural production.

There are three options to define how the model should consider the demand in services.

- **Exact:** demand and supply must meet precisely based on the iteration variables
- **Overshoot:** The indicated demand is the minimum requirement that should be met, and any excess is acceptable
- **Undershoot:** The indicated demand is the maximum limit that can be reached, and any shortfall is acceptable

When demand and supply must meet exactly, two options for iteration and convergence are available: either the convergence criteria is expressed as a percentage of the demand or as an absolute value in units of demand. Moreover, the first convergence criterion is the maximum average deviation between demanded and provided ecosystem goods and services for all demands together, and the second criterion is the maximum deviation for one type of demand. These values should be smaller than the amount of change between one year and the next, but greater than the amount of ecosystem goods and services produced by one cell. As an example, if a land use type has a demand of 2000 tons of crops and a cell production of 10 tons, the minimum value for the second convergence criterion should be (10/2000)*100% = 0.5%.

Table 4 Future demand for services under scenario "Specialisation" (left) and "Diversification" (right)

SPE	Built-up	AgriProd	AquaProd
2020	1682	43188	15489
2021	1697	43620	16883
2022	1712	44056	18403
2023	1728	44497	20059
2024	1743	44942	21864
2025	1759	45391	23832
2026	1775	45845	25977
2027	1791	46304	28315
2028	1807	46767	30863
2029	1823	47234	33641
2030	1840	47707	36668
Requirements	Exact	Undershoot	Undershoot

DIV	Built-up	AgriProd	AquaProd
2020	1682	32248	18564
2021	1697	33538	19864
2022	1712	34880	21254
2023	1728	36275	22742
2024	1743	37726	24334
2025	1759	39235	26038
2026	1775	40805	27860
2027	1791	42437	29810
2028	1807	44134	31897
2029	1823	45900	34130
2030	1840	47736	36519
Requirements	Undershoot	Undershoot	Undershoot

Conversion parameters

The simulations' temporal dynamics are determined by land use-specific conversions, which require three sets of parameters: conversion hierarchy, conversion resistance and land use transition sequences.

3.1 Conversion hierarchy

Changes in land use occur as a response to changes in demands. However, it is not immediately apparent how land uses will change in response to changes in demand, as each land use can provide multiple ecosystem goods or services and demand for these goods and services can be met by more than one land use. The conversion file provides information on the types of land changes that will occur to meet a particular demand. The values in the file are integers that reflect a hierarchy in the way land systems provide goods or services, with higher values indicating a greater level of provisioning. Negative values indicate land uses that are not expected to change in response to a particular demand. In order to meet an increase in demand, the model looks for locations with a land use that is indicated with a value greater than 0 in the conversion file and attempts to change this into another land use system with a higher value. The reverse is true for decreasing demands.

Table 5 Hierarchy of land system conversion as a response to changes in demands for each scenario (Specialisation on the left and Diversification on the right)

SPE	Built-up	AgriProd	AquaProd
TripleRice	-1	3	0
DoubleRice	-1	2	0
CashCrops	-1	1	0
Fruits	-1	2	0
AquaForest	-1	0	1
Aquaculture	-1	0	2
RiceAqua	-1	0	1
Urban	1	-1	-1

DIV	Built-up	AgriProd	AquaProd
TripleRice	-1	2	0
DoubleRice	-1	2	1
CashCrops	-1	1	0
Fruits	-1	3	0
AquaForest	-1	0	1
Aquaculture	-1	0	2
RiceAqua	-1	3	2
Urban	1	-1	-1

3.2 Conversion resistance

The first parameter set, conversion resistance, determines the ease of land use system change based on factors such as capital investment and demand. Each land use system must be assigned a value between 0 and 1, representing the relative resistance to change.

Table 6 Conversion resistance used for both scenarios

Conversion resistance	
TripleRice	0.4
DoubleRice	0.3
CashCrops	0.3
Fruits	0.6
AquaForest	0.8
Aquaculture	0.6
RiceAqua	0.2
Urban	1

3.3 Conversion matrix

The second set of parameters is the land use system-specific conversion settings, which are specified in a conversion matrix. This matrix outlines the allowed and disallowed land use system transitions, the regions where they can occur, and the time it takes for a transition to occur. The minimum and maximum number of years before a conversion can occur are indicated in the matrix, but the exact time depends on location-specific conditions and land use system pressure. The simulation of these interactions, combined with the constraints in the conversion matrix, determines the length of the period before a conversion occurs.

Table 7 Conversion matrix for the scenario "Specialisation"

SPE	TripleRice	DoubleRice	CashCrops	Fruits	AquaForest	Aquaculture	RiceAqua	Urban
TripleRice	Yes	No	No	Yes	No	No	No	Yes
DoubleRice	Yes	Yes	Yes	Yes	No	After 1 year	No	Yes
CashCrops	Yes	Yes	Yes	Yes	No	Yes	No	Yes
Fruits	No	No	Yes	Yes	No	No	No	Yes
AquaForest	No	No	No	Yes	Yes	No	No	No
Aquaculture	No	No	No	No	Yes	Yes	No	Yes
RiceAqua	No	No	No	No	No	Yes	No	Yes
Urban	No	No	No	No	No	No	No	Yes

Table 8 Conversion matrix for the scenario "Diversification"

DIV	TripleRice	DoubleRice	CashCrops	Fruits	AquaForest	Aquaculture	RiceAqua	Urban
TripleRice	Yes	Yes	Yes	Yes	No	No	Yes	Yes
DoubleRice	Yes	Yes	Yes	Yes	No	After 1 year	Yes	Yes
CashCrops	Yes	Yes	Yes	Yes	No	After 1 year	Yes	Yes
Fruits	After 1 year	After 1 year	After 1 year	Yes	No	After 1 year	After 1 year	Yes
AquaForest	No	No	No	Yes	Yes	After 1 year	After 1 year	No
Aquaculture	No	No	After 1 year	No	No	Yes	Yes	Yes
RiceAqua	After 1 year	Yes	Yes	After 1 year	After 1 year	After 1 year	Yes	Yes
Urban	No	No	No	No	No	No	No	Yes