Supplementary material

Supplement to: Soltani, Z., Abbasi-Shavazi, M.J., and Bagheri, A. (2024) Integrated Water-Population Interactions Framework: An Application to Assess Water Security in Iran. *Vienna Yearbook of Population Research*, 22. https://doi.org/10.1553/p-gjfn-7z5k

S1. Model description

To enhance comprehension of the application of the IWPI framework for the development WPSD model of Iran, this paper presents causal loop diagrams (CLDs) depicting the various subsystems of the model (for the population, employment, agriculture and water subsystem, respectively), along with a concise overview of the model's subsystems and its fundamental assumptions. The subsystems in the CLDs are explained below.

S1.1. Summary description of the population subsystem

The population subsystem was essential for estimating employment in the agricultural sector, water demand in agriculture and industry and food security. As shown in Figure S.1, the population subsystem considered the number of males and females of different ages, determined by survival and death indices and the fractional death rates among these age-sex groups. The total population was calculated based on the initial population, and births and deaths across different groups. The model also calculated the age dependency ratio, which interacts with employment and agricultural variables.

The model assumed that rural-urban migration rates increase when water resource limitations reduce rural income and increase rural unemployment, although other factors also influence these rates. Rural-urban migration linked the population and employment subsystems. Migrants were categorized into labor force, students and non-working/non-studying groups. Migration was driven by differences in urban and rural amenities, with agricultural livelihood insecurity (due to water shortages) increasing rural-urban migration by raising rural unemployment and reducing rural incomes.

In the model, it is assumed that changes in the number of immigrants affect the percentage of rural residents, thereby allowing the estimation of the urban or rural population. The estimation of changes in urban and rural household sizes is also used in calculating the number of urban and rural households. The model assumes that if the growth rate (obtained by subtracting the death rate from the fertility rate) becomes negative, the childbearing rates will increase.

S1.2. Summary description of the employment subsystem

The employment subsystem was the most complex aspect of the water security model, as depicted in Figure S2. This subsystem examined the impact of demographic changes on the labor force and how water resource limitations affect employment characteristics in urban and rural areas. The model included the effects of these changes on rural-urban migration rates and national food production patterns. Missing data, such as labor force leaving rates by sex, residence and education, were reconstructed using demographic techniques like active life expectancy by sex and educational level.

The model also reconstructed the educational status of the labor force population, considering factors such as entry into the workforce, illiteracy rates and dropout rates from different Educational levels. Those leaving education entered the labor market based on participation rates. According to Iranian data, it takes at least one year for a person who has graduated from any level of education to get a job, meaning that there is a one-year delay between graduation and employment. Additionally, a one-year break has been factored in for workers changing their job group. The labor force was divided into unemployed and employed populations across the industry, agriculture and services sectors. Employment growth in each sector was constrained by growth capacity, influenced by water insecurity.

The active population could be employed in any economic subgroup of industry, agriculture and services, depending on the capacity of that economic group, based on its sex composition and capacity level, and whether it was urban or rural. The expected income in each job group also influenced the priority of recruitment. The model assumed that changes in the age dependency ratio affected both the unemployment rate and the participation rate, and thus influenced active life expectancy.

S1.3. Summary description of the agricultural subsystem

The agricultural subsystem addressed two key aspects of water security: meeting the food needs of the population and ensuring the livelihood of the agricultural community (Figure S.3). Food security, in this context, was determined by the food requirements of different age and sex groups, and was linked to water security at a macro level. Micro-level factors like household income and financial capability were not considered. However, agricultural waste and losses were included in the food security assessment.

Food security was calculated based on the difference between agricultural production (across 11 main food groups) and the gross desired food needs of the population, which included net desired food needs and agricultural waste and losses. It was assumed that food insecurity for any agricultural product group could be mitigated through imports, although high import costs could hinder investments in reducing agricultural losses, thus maintaining a high gross need for agricultural products despite these policies.

The model incorporated variables affecting agricultural water use to estimate agricultural production (crops, horticulture and livestock). The agricultural population was modeled based on the rural population, household size, livelihood composition and age dependency ratio. The number of farmers, gardeners and livestock farmers was derived from the agricultural population. Livestock and horticultural production were assumed to follow past trends and population food needs, influenced by the number of livestock farmers and gardeners and feed crop production.

Agricultural production depended on the number of farmers, land area, yield per hectare (influenced by educational levels) and cultivation patterns. Cultivation patterns could be adjusted to reduce water consumption, increase food self-sufficiency and enhance livelihood security.

In terms of agricultural livelihoods, the model used an index of minimum area to support livelihood security. Comparing this index with the land per capita owned by farmers, alongside rural household size, provided an estimate of the population exposed to livelihood insecurity. An increase in population exposed to agricultural livelihood insecurity was expected to drive rural-urban migration, as discussed in the employment subsystem.

S1.4. Summary description of the water subsystem

The water subsystem, described in Figure S.4, was comprised of three main components: modeling of water resources, modeling of water consumption and assessment of the population exposed to water insecurity. Water resource modeling employs common water balance models that separately consider surface and groundwater resources.

While dynamic system models typically limit water demand growth through pricing, for water consumption modeling, this research did not include water pricing due to the country's subsidized water and energy policies. Domestic water demand was modeled as a function of per capita consumption and population size, which increases as the population grows. An increase in domestic water demand beyond allowable levels resulted in more people being exposed to water insecurity, and reduced allowable agricultural water over time. Per capita domestic water consumption was also linked to household size.

Industrial water consumption modeling considered population growth, economic activity and precipitation and evaporation changes, although climate change effects were excluded. Future industrial water demand was estimated based on the urban population size, the livelihood composition and the age dependency ratio affecting the labor force. Differences between allowable industrial water and demand indicated industrial water scarcity, which were used to calculate the population exposed to industrial water insecurity. Long-term increasing levels of industrial water insecurity could raise the allocation of water to this sector.

Agricultural water consumption was linked to agricultural production through the water footprint of agricultural products. Comparing annual agricultural water consumption with available water indicated that there is pressure on water resources. If annual water needs are unmet by renewable resources, groundwater reserves are withdrawn, leading to long-term depletion and reduced extraction capability, threatening food security and agricultural livelihoods.

Water security was measured by the population exposed to water insecurity, including those at risk of domestic and industrial water insecurity due to resource limitations and those at risk of food insecurity and agricultural livelihood insecurity, influenced by factors such as cultivation patterns and land availability. Current Iranian water resource management policies prioritize drinking and industrial water over agricultural water. Thus, the model assumed that growing demand and increasing water insecurity in the domestic and industrial sectors will eventually lead to a reallocation of allowable water, reducing agricultural water allocation. This insufficiency in water resources for all sectors could further reduce water reserves and agricultural water allocation.

https://doi.org/10.1553/p-gjfn-7z5k

Figure S.1 The causal loop diagram of the population subsystem

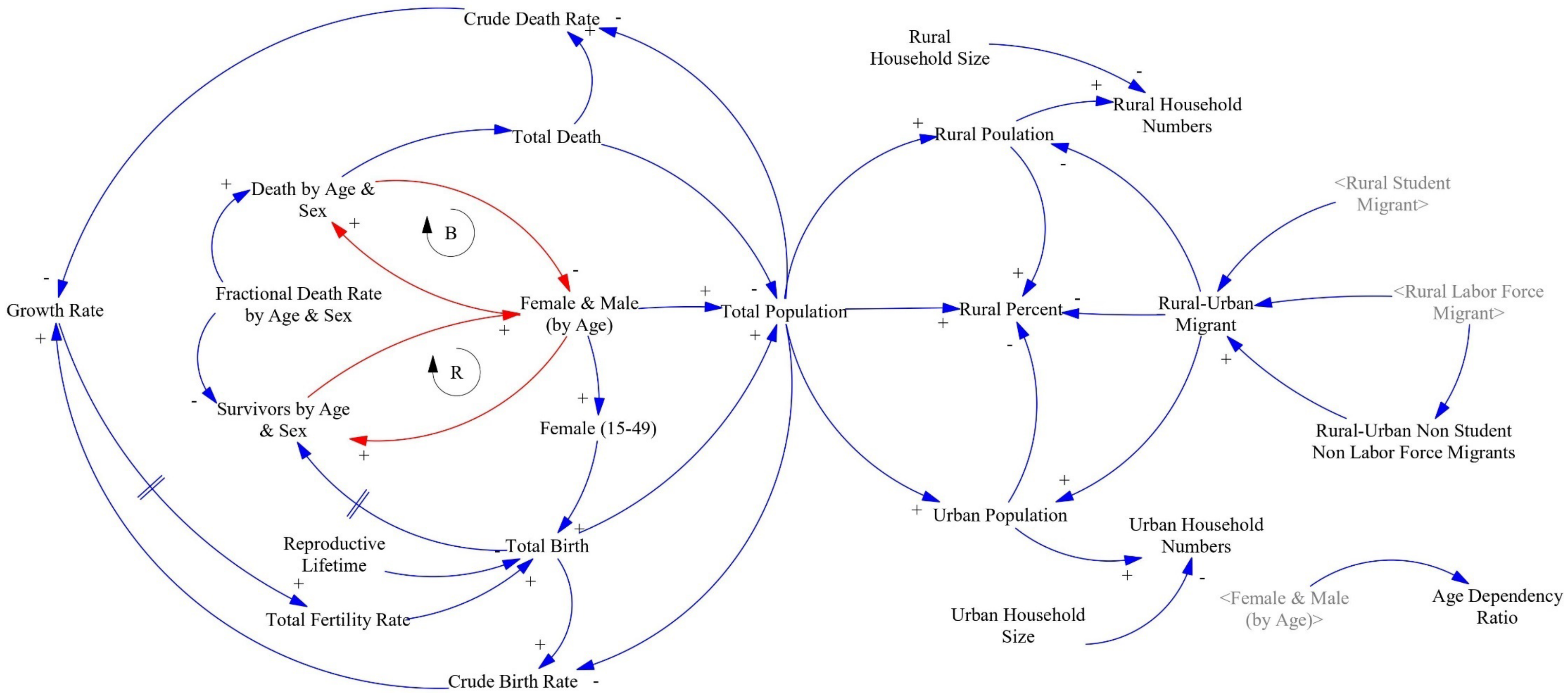


Figure S.2 The causal loop diagram of the employment subsystem

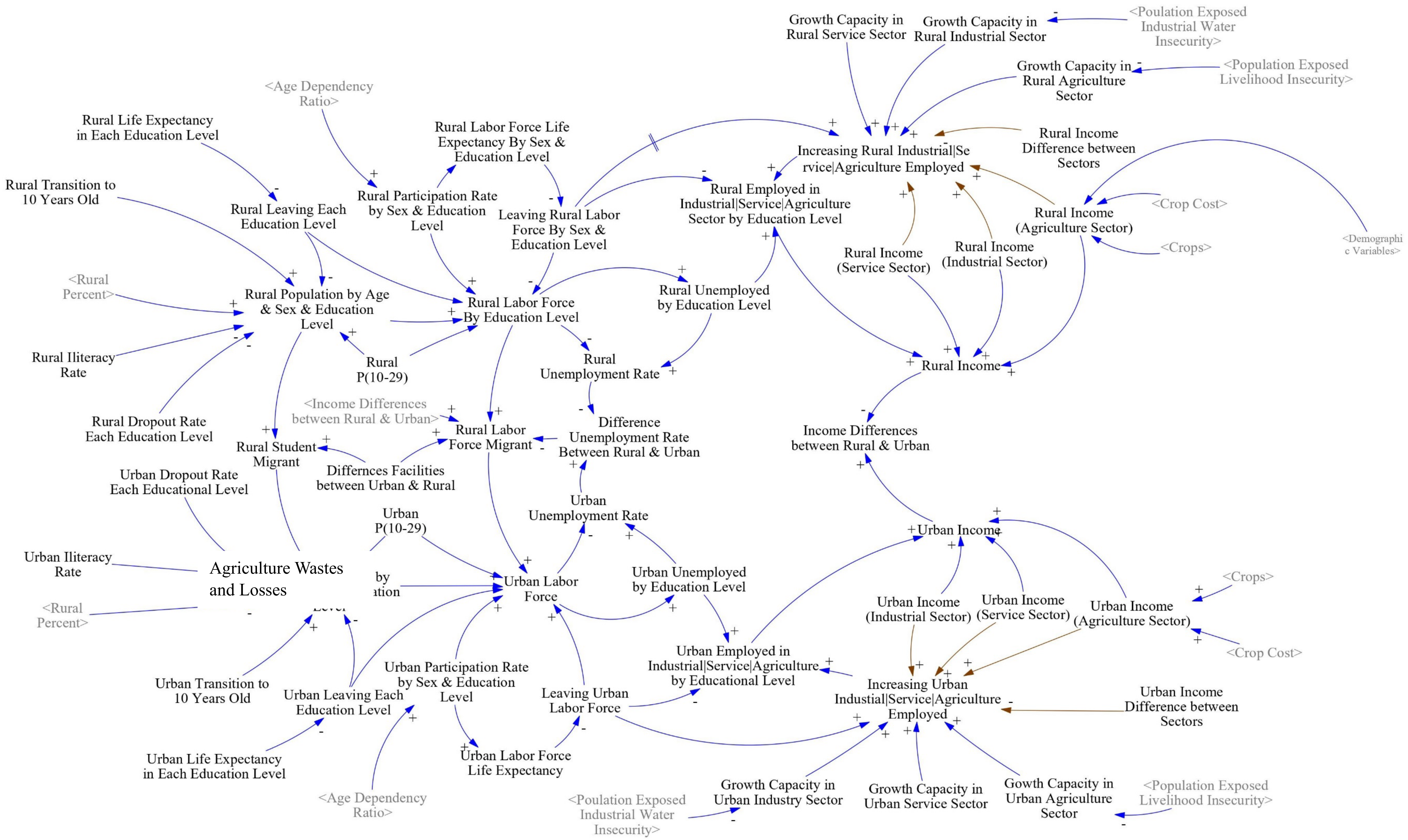


Figure S.3 The causal loop diagram of the agricultural subsystem

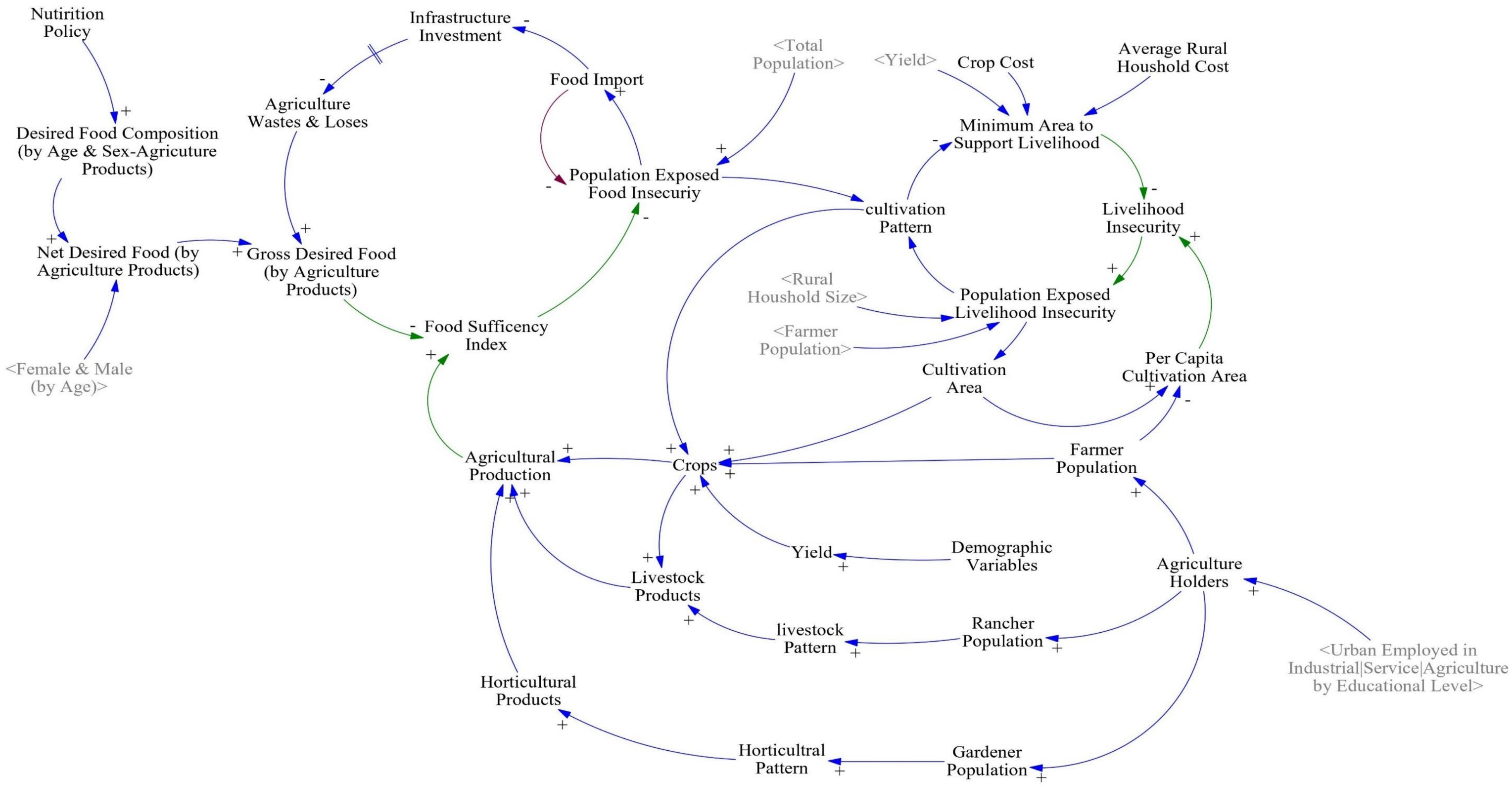
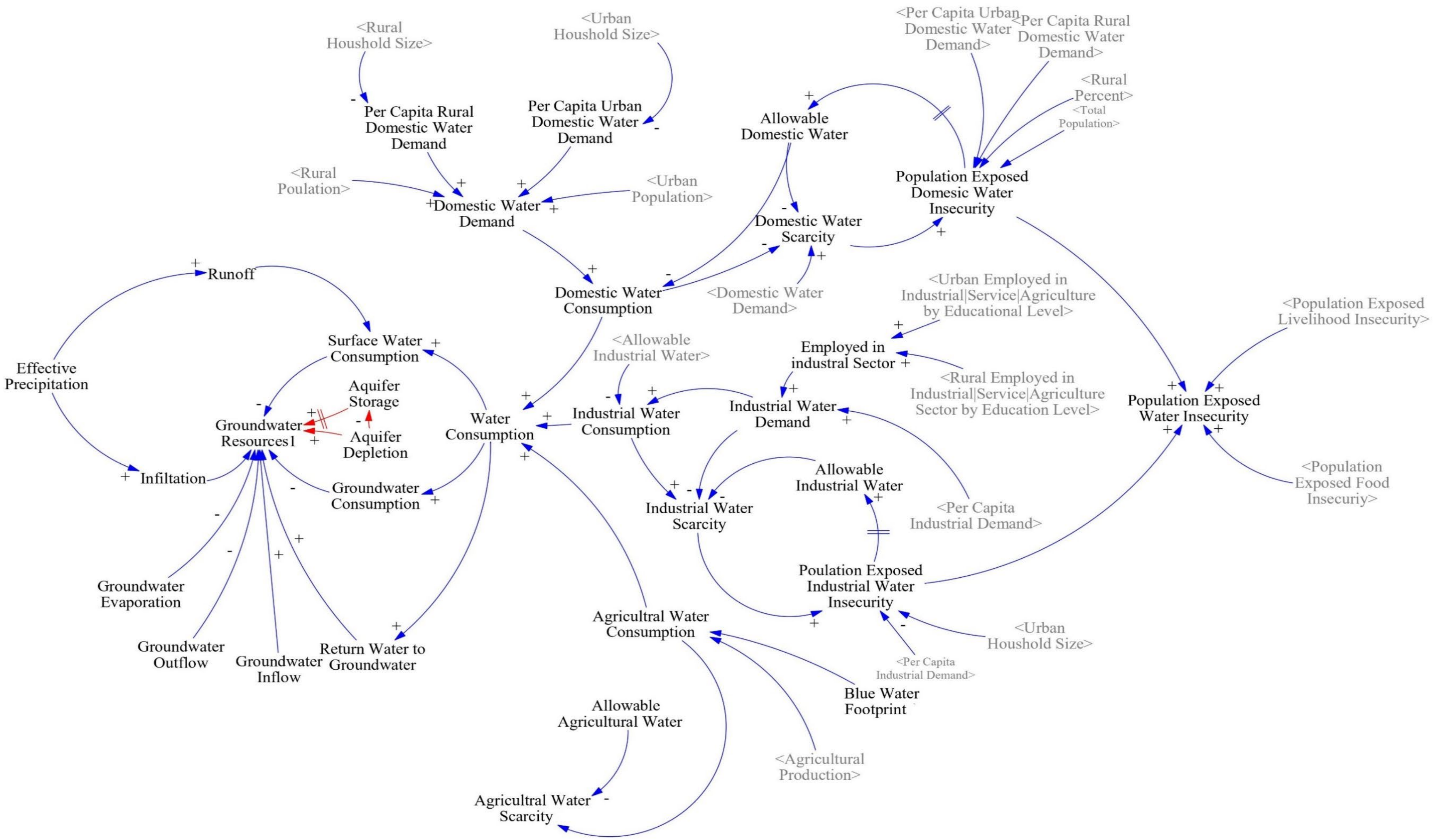


Figure S.4 The causal loop diagram of the water subsystem



S2. Mathematical formulation of the model

Some of key equations used in the WPSD model are as follows:
According to the groups based on food security, 12 main groups of food crops were differentiated in the modeling as shown in Table S1.

Table S. 1 The main groups of food crops used in the model

i	Food group
1	Cereals (excluding rice)
2	Rice
3	Beans
4	Potatoes
5	Vegetables
6	Oilseeds
7	Sugar
8	Fruits
9	Red meat
10	White meat
11	Eggs
12	Milk

Equation S1 shows the sufficiency of production of each food product, *Suf(i)*, which is the ratio of the production of product, *Prod(i),* and the gross requirement of food product, *GDC(i)*, in the country (meaning the food requirement of the population, including the waste and losses of the agricultural production). If the production exceeds the gross requirement of the product, then the sufficiency of the production of the product is considered to equal 1.

$$Suf(i) = \begin{cases} \frac{Prod\ (i)}{GDC\ (i)} & if\ \ Prod(i) < GDC(i) \\ 1 & if\ Prod(i) \geq GDC(i) \end{cases}, \quad i = 1 \ldots 12 \tag{S.1}$$

where
Suf: Sufficiency of product
Prod: Production of product
GDC: Gross desired consumption of product

The priority for cultivation of each product focused on food security (self-sufficiency, *CPF*) is determined by the percentage of each agricultural product's sufficiency to meet the total food needs of the population, as outlined in equation S2:

$$CPF\ (i) = \begin{cases} 1 & if\ Suf(\,i) < 20\% \\ 2 & if\ \ 20\% \leq Suf(\,i) < 30\% \\ 3 & if\ \ 30\% \leq Suf(\,i) < 40\% \\ 4 & if\ \ 40\% \leq Suf(\,i) < 50\% \\ 5 & if\ \ 50\% \leq Suf(\,i) < 60\% \\ 6 & if\ \ 60\% \leq Suf(\,i) < 70\% \\ 7 & if\ \ 70\% \leq Suf(\,i) < 80\% \\ 8 & if\ \ 80\% \leq Suf(\,i) < 90\% \\ 9 & if\ \ 90\% \leq Suf(\,i) < 100\% \\ 10 & if\ Suf(\,i) \geq 100\% \end{cases}, \qquad i = 1 \dots 12 \tag{S.2}$$

where
Suf: Sufficiency of product
CPF: Cultivate priority based on food security

As shown in equation S3, the food product sufficiency index, *FSI*, is estimated based on the average food product adequacy of each food product. Of course, this index is calculated in a different way in the model in different food policies, with other acceptable self-sufficiency percentages.

$$FSI = \sum_{i=1}^{12} Suf(i)/12, \tag{S.3}$$

where
FSI: Food sufficiency index
Suf: Sufficiency of products

Equation S4 determines the minimum per capita area for crop *i* available to each rural household, *MAL*, to ensure that the family's income from crop *i* suffices to cover the average expenses of the rural household.

$$MAL(i) = \frac{ACRH}{Income(i)} \times Area(i), \qquad i = 1 \dots 12 \tag{S.4}$$

where
MAL: Minimum area for livelihood security
ACRH: Average cost of rural household
Income: Income from crop in per hectare
Area: Crop area (hectare)

The priority for cultivation aimed at providing agricultural livelihood is calculated according to equation S5, based on the index of the minimum area required for each crop *i* to sustain a rural household. Thus, the less land area needed to provide livelihood security, the higher the priority for cultivating that product.

$$CPL\ (i) = \begin{cases} 1 & if\ MAL(i) \leq 2 \\ 2 & if\ 2 < MAL(i) \leq 3 \\ 3 & if\ 3 < MAL(i) \leq 4 \\ 4 & if\ 4 < MAL(i) \leq 5 \\ 5 & if\ 5 < MAL(i) \leq 6 \\ 6 & if\ 6 < MAL(i) \leq 7 \\ 7 & if\ 7 < MAL(i) \leq 8 \\ 8 & if\ 8 < MAL(i) \leq 9 \\ 9 & if\ 9 < MAL(i) \leq 10 \\ 10 & if\ MAL(i) > 10 \end{cases}, \qquad i = 1 \dots 12 \tag{S.5}$$

where

https://doi.org/10.1553/p-gjfn-7z5k

CPL: Cultivate priority based on livelihood security
MAL: Minimum area for livelihood security

The cultivate priority of each agricultural product CP(i) is obtained based on the average of priority of finding food security and the priority of finding livelihood security, according to equation S6.

$$CP\ (i) = \frac{[(CPF(i) \times FP) + (CPL(i) \times LP)]}{(FP + LP)}, \qquad i = 1 \ldots 12 \tag{S.6}$$

where
CP: Cultivate priority of product
CPF: Cultivate priority of product based on food security
CPL: Cultivate priority of product based on livelihood security
FP: Food security
LP: Livelihood security

The total population that is exposed to food insecurity, *PEFI*, is estimated based on the index of food sufficiency by equation S7.

$$PEFI = TP \times (1 - FSI) \tag{S.7}$$

where
PEFI: Population exposed to food insecurity
TP: Total population
FSI: Food sufficiency index

Since the lack of adequate income impacts both farmers and their households, equation S8 calculates the number of people exposed to agricultural livelihood insecurity, *PEILA*, by estimating the number of people whose livelihoods depend directly or indirectly on agriculture and who have less than the minimum required land area for household sustenance. This formula assumes agricultural activity is the sole source of family livelihood.

$$PEILA = \frac{(FMALP) \times (UHS \times UEA + REA \times RHS)}{(UEA + REA)}, \tag{S.8}$$

where
PEILA: Population exposed to agricultural livelihood insecurity
FMALP: The percentage of farmers with area less than the average minimum area for livelihood security
UHS: Urban household size
UEA: Urban employed in agricultural sector
RHS: Rural household size
REA: Rural employed in agriculture sector

The population exposed to livelihood insecurity in the industrial sector, *PEILI,* is calculated by a combination of the estimation of water scarcity in the industrial sector, the number of employees in the industrial sector and the average size of the household, according to equation S9.

$$PEILI = IEEIWS \times (UEI \times UHS + REI \times RHS)/(UEI + REI) \quad (S.9)$$

where
PEILI: Population exposed to industrial livelihood insecurity due to water shortage
IEEIWS: Industrial employees exposed industrial water scarcity
UEI: Urban employed in industrial sector
UHS: Urban household size
REI: Rural employed in industrial sector
RHS: Rural household size

The population that is exposed to domestic water insecurity is calculated by equation S10, based on the comparison with domestic allowable (allocable) water.

$$\mathrm{PEDWI} = \begin{cases} 0 & if\ \mathrm{DWD} \leq \mathrm{DAW} \\ \dfrac{(\mathrm{DWD} - \mathrm{DAW})}{\mathrm{TCDWW}} & if\ \mathrm{DWD} > \mathrm{DAW} \end{cases} \quad (S.10)$$

where
PEDWI: Population exposed to domestic water insecurity
DWD: Domestic water demand
DAW: Domestic allowable water
TCDWW: Total per capita domestic water withdrawal

S3. Comparison of Model results to historical data

In order to compare the model outputs with historical data trends, the alterations in key variables were accurately investigated across four distinct subsystems. The examination of population and employment variables involved a comparative analysis with the outcomes of Iran's last population and housing census in 2016. For the agriculture and water subsystems, the focus was on contrasting changes in key variables with data published by the relevant ministries over an eight-year period following the start of the simulation.

Within the **population subsystem**, Table S.2 presents a comparison of key variables with observed data. The model demonstrates a commendable estimation accuracy, capturing 95% of demographic variables during the initial five-year period of simulation. Nevertheless, discrepancies arise in the model's reconstruction of other population variables.

Table S.2 Comparison of key variables of the population subsystem to Census data in 2016 (five years after the starting point of the simulation)

Variable	Total population (millions)	Total households (millions)	Rural-Urban migrants	Rurality ratio	Sex ratio	Age dependency ratio	Urban household size	Rural household size
Model results in 2016	79.62	23.54	643 358	0.27	1.01	0.43	3.31	3.6
Census data in 2016	79.90	24.19	525 116	0.24	1.03	0.43	3.30	3.4
Relative percentage error	0.35%	2.69%	22.52%	13.98%	1.75%	0.06%	0.24%	6.1%

https://doi.org/10.1553/p-gjfn-7z5k

The **employment subsystem** delves into the reconstruction of employment data based on gender and educational levels within the active population. To reconstruct the data of employees, we initially focused on reconstructing the student population within various study groups. Table S.3 presents the results of our comparative reconstruction of the student population across different educational levels, delineated by gender and place of residence (urban or rural). Our findings suggest that the model was more accurate in reconstructing the student population in urban areas than in rural areas. This increased accuracy in urban areas can likely be attributed to these areas having lower educational disparities and dropout rates than rural areas. However, it is noteworthy that the reconstruction error generally remained below 10% across all groups, with the exception of rural female elementary students.

Table S.3 Comparison of the number of students enrolled to census data in 2016 (five years after the starting point of the simulation), by residential status, gender and educational level

Source	Residential status	Gender	Educational level		
			Elementary	High school	Higher education
Model results for 2016	Urban	Male	1 080 516	1 945 762	2 336 219
		Female	998 786	1 783 589	2 099 180
	Rural	Male	473 201	669 099	379 268
		Female	427 651	531 630	257 657
Census 2016	Urban	Male	1 081 000	1 937 000	2 239 000
		Female	1 042 000	1 212 000	2 041 000
	Rural	Male	446 500	612 100	385 500
		Female	368 800	523 700	244 500
Relative percentage error	Urban	Male	0.04%	-0.45%	-4.34%
		Female	4.15%	6.72%	-2.85%
	Rural	Male	-5.98%	-9.31%	1.62%
		Female	-15.96%	-1.51%	-5.38%

Subsequently, we reconstructed the data related to employees based on their entry into the post-education labor market. In Table S.4, the model accuracy was higher in reconstructing groups with larger populations, whereas the accuracy was diminished in reconstructing smaller population groups, such as the urban illiterate workforce and the rural population with higher education. Nonetheless, the model exhibited particularly high accuracy in reconstructing the active population with higher education in urban areas, and the population with secondary and primary education in rural areas. This can be attributed to the lower likelihood of activity gaps due to reasons such as returning to education in these main groups.

Table S.4 Comparison of the number of population in the employment subsystem to Census data in 2016 (5 years after the starting point of the simulation), by residential status and educational attainment

Source	Residential status	Educational level			
		Illiterate	Elementary	High school	Higher education
Model results for 2016	Urban	1 152 917	2 559 887	9 455 565	5 853 541
	Rural	1 258 022	1 845 485	3 079 611	637 322
Census 2016	Urban	931 600	2 298 800	8 411 100	5 753 000
	Rural	1 392 300	1 877 300	3 014 200	825 900
Relative percentage error	Urban	-23.76%	-11.36%	-12.42%	-1.75%
	Rural	9.64%	1.69%	-2.17%	22.83%

Table S.5 Comparison of the agricultural products in the agriculture subsystem to real data from Ministry of Agriculture-Jihad (2011 to 2018)

Agricultural products	Production (in thousand tons)	2011	2012	2013	2014	2015	2016	2017	2018
Cereals (excluding rice)	Observed	13.12	13.38	13.97	15.2	15.89	19.49	16.44	17.35
	Model result	13.11	11.84	12.97	13.59	14.01	14.28	14.5	14.63
	Relative percentage error	0%	-13%	-8%	-12%	-13%	-36%	-13%	-19%
Rice	Observed	1.89	2.36	2.45	2.35	2.35	2.92	3.21	3.11
	Model result	1.89	2.45	2.57	2.69	2.81	2.93	3.05	3.16
	Relative percentage error	0%	4%	5%	13%	16%	0%	-5%	2%
Beans	Observed	0.42	0.46	0.5	0.62	0.52	0.67	0.7	0.75
	Model result	0.42	0.46	0.51	0.53	0.56	0.58	0.6	0.65
	Relative percentage error	0%	0%	2%	-17%	7%	-16%	-17%	-15%
Potatoes	Observed	4.71	5.07	4.6	4.99	5.14	5	5.02	5.14
	Model result	4.71	5.43	5.93	6.43	6.93	7.44	7.95	8.47
	Relative percentage error	0%	7%	29%	29%	35%	49%	58%	65%
Vegetables	Observed	14.03	13.44	14.05	13.48	15.48	15.67	15.44	16.4
	Model result	11.2	11.44	12.51	13.59	14.7	15.82	16.95	18.1
	Relative percentage error	-20%	-15%	-11%	1%	-5%	1%	10%	10%
Fruits	Observed	19.1	19.59	21.19	22.08	22.86	24.64	24.64	23.66
	Model result	22.53	23	24.7	26.4	28.12	29.8	31.5	33.2
	Relative percentage error	18%	17%	17%	20%	23%	21%	28%	40%
Oilseeds	Observed	0.31	0.31	0.31	0.22	0.16	0.21	0.23	0.31
	Model result	0.42	0.45	0.44	0.44	0.43	0.42	0.4	0.39
	Relative percentage error	35%	45%	42%	100%	169%	100%	74%	26%
Sugar	Observed	10.35	9.43	10	11.32	13	13.45	15.88	12.49
	Model result	10.35	10.01	10.91	11.83	12.74	13.66	14.57	15.47
	Relative percentage error	0%	6%	9%	5%	-2%	2%	-8%	24%
Red meat	Observed	0.74	0.75	0.76	0.79	0.81	0.82	0.84	0.83
	Model result	0.74	0.8	0.81	0.84	0.85	0.86	0.87	0.88
	Relative percentage error	0%	7%	7%	6%	5%	5%	4%	6%

https://doi.org/10.1553/p-gjfn-7z5k

Agricultural products	Production (in thousand tons)	2011	2012	2013	2014	2015	2016	2017	2018
White meat	Observed	2.44	2.61	2.85	2.98	3.11	3.16	3.44	3.62
	Model result	2.43	2.51	2.64	2.77	2.91	3.04	3.19	3.33
	Relative percentage error	0%	-4%	-7%	-7%	-6%	-4%	-7%	-8%
Eggs	Observed	0.7	0.91	0.89	0.93	0.93	0.94	0.89	0.9
	Model result	0.7	0.79	0.81	0.83	0.85	0.87	0.89	0.91
	Relative percentage error	0%	-13%	-9%	-11%	-9%	-7%	0%	1%
Milk	Observed	7.69	7.95	8.27	8.8	9.14	9.65	10.18	10.59
	Model result	7.69	8.16	8.43	8.71	8.99	9.28	9.57	9.87
	Relative percentage error	0%	3%	2%	-1%	-2%	-4%	-6%	-7%
Industrial plants	Observed	0.29	0.23	0.21	0.2	0.2	0.18	0.2	0.18
	Model result	0.29	0.13	0.13	0.13	0.13	0.13	0.13	0.13
	Relative percentage error	0%	-43%	-38%	-35%	-35%	-28%	-35%	-28%
Fodder plants	Observed	14.82	14.95	15.53	18.44	19.7	20.7	20.18	20.95
	Model result	14.82	15.45	16.16	16.87	17.57	18.27	18.96	19.46
	Relative percentage error	0%	3%	4%	-9%	-11%	-12%	-6%	-7%

Within the agricultural subsystem, as illustrated in Table S.5, production changes of 14 agricultural product groups were reviewed against agricultural yearbooks up to eight years after the start of the simulation. While the model generally aligns with actual trends, discernible cross-sectional distortions in the evolution of agricultural products suggest limitations in its predictive capabilities (especially for products such as potatoes, fruits, oilseeds and industrial plants).

Turning to the **water subsystem** in Table S.6, we can see that the model's simulation of aquifer decline exhibits a trajectory closely mirroring real data for the initial eight years. However, a steeper decline in simulated aquifer levels, compared to actual data, is observed. This disparity may be attributed to the model's failure to incorporate the influence of climate change, particularly the seasonal increase in rainfall, impacting aquifer pressure, which remains unaccounted for in the model (due to privacy concerns regarding data on Iran's water situation, we have excluded the reports about the decline of Iran's aquifer level. We apologize for any confusion this may cause, and will only present the corresponding figures here).

Table S.6 Comparison of the predicted drop in Iran's water aquifer level to the actual drop in the water aquifer level according to the statistics of the Iran Water Resources Management Company

Year	2011	2012	2013	2014	2015	2016	2017	2018
Relative percentage error	0%	0%	0%	0%	1%	1%	-2%	-12%

https://doi.org/10.1553/p-gjfn-7z5k