We deeply appreciate the detailed and constructive comments provided by the three anonymous reviewers. Following their suggestions and comments, we have extensively revised the manuscript and provided a point-to-point response to each comment. The original comments are in **bold** font, our response is in regular font, and the changes in the text are in blue.

Comment 1

This paper introduces CHANS-SD-YRB v1.0, a system dynamics CHANS model of the Yellow River Basin and operates at provincial (human processes) and sub-basin (natural processes) scales. The manuscript reads well and a supplement with more elaborate model descriptions. By coupling the human and natural processes within a regional CHANS framework, this model description paper lays the foundations for policy-oriented modelling for the Yellow River Basin. Overall, the model is an ambitious synthesis, but several methodological choices and transparency gaps need some tightening. Find my specific comments and suggestions below.

Response: Thank you for taking your time to carefully review our study and providing insightful and constructive comments. Following the suggestions and comments, we have extensively revised the manuscript.

1. Introduction

The first two paragraphs of the introduction goes back and forth between regional and global CHANS research and modelling, which breaks the flow and is a little disorientating. It would be clearer if CHANS research and modelling is first introduced more generally. After introducing the main models, I suggest that the authors identify the overrepresentation of global scale CHANS models. And then bring in the distinction and need for more regional scale modelling, which situates this work within.

Response： Thank you for the comments on organizing the introduction. We have reorganized the first two paragraphs following suggestions. The revised texts are shown below:

Coupled human and natural systems (CHANS) emphasize the reciprocal feedback and co-evolution between human and natural systems, offering an integrated framework for diagnosing complex problems and guiding sustainable development (Fu and Li, 2016). By capturing dynamic interactions of interconnected components, CHANS theories and models enable more effective policies and interventions that align ecological integrity with socioeconomic progress (Motesharrei et al., 2016; Verburg et al., 2016). Numerous integrated modelling approaches have been developed to simulate human–natural interactions at the global scale. These include system dynamics-based integrated assessment models (IAMs), such as ANEMI (Breach and Simonovic, 2021), FeliX (Rydzak et al., 2013; Ye et al., 2024), and FRIDA (Rajah et al., 2025), process-based and optimization-based IAMs (Vaidyanathan, 2021), and Earth system models with synchronously coupled human components, such as E3SM-GCAM (Di Vittorio et al., 2025) and integrated Earth system models (iESMs) (Jain et al., 2022). These models effectively characterize human–natural interactions at the global scale, and have been applied to assess the impacts of climate change on human society (e.g., agriculture (Monier et al., 2018), economic damage (Wang et al., 2020), fatalities increase and welfare loss (Dottori et al., 2018)), and humans' feedback on the Earth system, such as those from climate mitigation on water and food security (Cheng et al., 2022; Fujimori et al., 2022).

Currently, most CHANS models are at the global scale (Calvin and Bond-Lamberty, 2018) with much fewer regional models. While global modeling research has deepened our knowledge of dynamic feedbacks among Earth's spheres and system evolution under climate change, sustainability challenges often manifest at regional scales, where social and ecological dynamics are more intricately intertwined (Liu et al., 2007). Compared to global CHANS, regional CHANS are open systems that continuously exchange energy and materials with other regions and the global system, e.g., water resources and electricity transfers (Dobbs et al., 2023; Zhang et al., 2022) and trade (Ristaino et al., 2021), resulting in pericoupling and telecoupling of different systems (Liu, 2017). Besides, regional CHANS are shaped by more immediate and complex human influences and stressors, such as urban expansion (van Vliet, 2019), ecological protection (Xu et al., 2017; Yang et al., 2022), and water resource regulation policies (diversion and allocation) (Song et al., 2024), which alter regional CHANS dynamics. Due to their diverse ecological and socioeconomic resilience, regional CHANS exhibit heterogeneous responses to external weather events and climate change, as evidenced by differing responses in crop yield (Hasegawa et al., 2021) and economic production to extreme heat and warming (Waidelich et al., 2024a). Furthermore, the coarse spatiotemporal resolution of global models limits their capacity to support effective decision-making for regional development (X. Li et al., 2018). As such, advancing regional CHANS modeling is essential for informing adaptive strategies in the face of growing regional environmental and societal pressures.

L66-69: The sentence on CHANS theories and models applies to both regional and global scale. I suggest moving this above, prior to the delineation between global vs. regional CHANS models for a better flow.

Response： Thanks! This is a very good suggestion. We moved the CHANS theory and method sentences before the introduction of the global model. The revised paragraph reads much better:

Coupled human and natural systems (CHANS) emphasize the reciprocal feedback and co-evolution between human and natural systems, offering an integrated framework for diagnosing complex problems and guiding sustainable development (Fu and Li, 2016). By capturing dynamic interactions of interconnected components, CHANS theories and models enable more effective policies and interventions that align ecological integrity with socioeconomic progress (Motesharrei et al., 2016; Verburg et al., 2016). Numerous integrated modelling approaches have been developed to simulate human–natural interactions at the global scale.

L70-71: I suggest rephrasing it as "growing regional environmental and societal pressures"

Response: This is excellent suggestion! Thank you for the comments.

We have rephrased it. The revised texts are shown below:

As such, advancing regional CHANS modeling is essential for informing adaptive strategies in the face of growing regional environmental and societal pressures.

L75-81: The sentence mixes model classes (e.g. IAMs, iESMs) and specific methodologies (e.g., system dynamics), which makes it slightly misleading – since the SD models can also be considered IAMs. I recommend revising this sentence to clearly distinguish model classes from methodologies and examples.

Response: We sincerely appreciate your suggestions and the specific examples of revisions provided. We have incorporated these valuable suggestions into the revised manuscript to improve the clarity of our argument.

The revised texts are shown below:

These include system dynamics-based integrated assessment models (IAMs), such as ANEMI (Breach and Simonovic, 2021), FeliX (Rydzak et al., 2013; Ye et al., 2024), and FRIDA (Rajah et al., 2025), process-based and optimization-based IAMs (Vaidyanathan, 2021), and Earth system models with synchronously coupled human components, such as E3SM-GCAM (Di Vittorio et al., 2025) and integrated Earth system models (iESMs) (Jain et al., 2022).

2. Model description

L162: The citation provided for the sentence is a model description paper of a specific SD model. Given that the sentence is on SD method more generally, I would suggest references to more foundational works (e.g., Forrester, 1968; Richardson, 2011)

1. Forrester, J. W.: Principles of systems, Pegasus Communications, Inc., Waltham, MA, 1968.

2. Richardson, G. P.: Reflections on the foundations of system dynamics: Foundations of System Dynamics, Syst. Dyn. Rev., 27, 219–243, https://doi.org/10.1002/sdr.462, 2011.

Response: Thank you for recommending these classic works. We have already added them as references in our manuscript.

The revised texts are shown below:

We developed the CHANS-SD-YRB based on system dynamics, a method well-suited for capturing complex system behaviors characterized by nonlinearity, multi-level structures, and feedback loops (Forrester, 1968; Richardson, 2011).

2.2.1 Population

In Figure 3, the only direct cross-sectoral feedback is between the Economy and Population, through life expectancy. Total fertility is kept exogenous to the model, yet in most SD models fertility rate tends to be endogenized with a function from GDP per capita – assumption being that fertility rate is negatively correlated with higher literacy levels and access to contraceptives etc. What is the justification for excluding this relationship in the model?

Response: We appreciate the reviewer for highlighting this critical issue. We have taken this opportunity to provide further clarification regarding the fertility rate setting.

We agree that in general SD models, fertility rates are often endogenized as a function of GDP per capita, reflecting the declining fertility rate following economy growth and education improvement.

However, in China, government policy has historically played a dominant role in regulating trajectory of population growth. The strict implementation of the Family Planning Policy (particularly the One-Child Policy) significantly altered fertility trends, making them deviating from those of economic indicators than in other countries. In this specific context, modeling fertility solely based on GDP would fail to capture these strong policy-driven constraints. Therefore, to reflect the policy influence and ensure the accuracy of the simulation, we treated the Total Fertility Rate (TFR) as an exogenous variable based on historical data and policy targets, rather than endogenizing it through economic feedback loops. By designing a scenario where fertility rates respond dynamically to economic conditions, the model can evaluate the impact of family planning policies.

We have clarified this justification in Section 2.2.1 of the revised manuscript. The revised texts are shown below:

Population dynamics are driven by births, deaths, and migration (Fig. 3). Births are calculated based on exogenous total fertility rates derived from historical data (Equation S2) to reflect the strong influence of China's Family Planning Policy.

Migration rate is also depicted as an exogenous input to the model. In the model, this appears to be an endogenous variable as a function of GDP difference between national and YSB, which scales the exogenous parameter, migration rate other effects. Also, the migration rate other effects is a time-dependent parameter (lookup table) that increases migration slightly between 2000 and 2005, sharp rise by 2010, and then a decline by 2015. Thereafter, this parameter is held constant. How was this determined? How does this affect future projections of migration? The modelling choices and assumptions here should be clearly documented in the model description.

Response: Thank you for pointing out this important issue in the model.

The reviewer is correct that migration is modeled as an endogenous variable driven by GDP disparities; however, an exogenous parameter, "Migration Rate Other Effects," was introduced to calibrate the model against historical data to account for the effect of other non-economic factors (e.g., policy changes or social networks) not explicitly captured by the GDP function. This approach ensures that the model reproduces historical trends accurately. Since actual migration data is reported every five years (starting from 2000), this calibration factor is for each five-year, and VENSIM automatically applies linear interpolation to generate a continuous yearly time series.

This parameter is held constant after 2015, assuming that the magnitude of non-economic effects on migration remains at the most recent level, preventing arbitrary fluctuations in future projections.

Since there are many similar parameters in the model, for the future scenario, we clarified in section 3.2.1 that with the exception of variables specifically adjusted for future scenario settings (file exogenous time-varying inputs.xlsx), all other parameters are held constant at their most recent historical levels as:

For future projections, variables or parameters not specified in the table are held constant at historical levels from their most recent year.

L237-238: Births, deaths, and migration are "determined by a series of social and economic factors" is very vague. What and how are these factors conceptualised and represented in the model? Given that life expectancy is endogenous and part of the key inter-sectoral feedback, I would expect more descriptions of how life expectancy is modelled.

Response: Thank you for the comments.

We apologize for vague description in the original manuscript without referencing to the detailed mathematical descriptions for birth, death, and migration in the Supporting Information Section S2.1.

In the population sector, births are derived from exogenous fertility rates. Regarding life expectancy, it is endogenously modeled based on the Human Well-being Index. This index serves as a proxy for living standards and is computed using GDP per capita and years of schooling. Migration is driven by economic disparities represented by GDP per capita between the YRB and the national average level.

In the revised manuscript, we have explained the specific mechanisms driving population dynamics to improve clarity. The revised texts are shown below:

Population dynamics are driven by births, deaths, and migration (Fig. 3). Births are calculated based on exogenous total fertility rates derived from historical data (Equation S2) to reflect the strong influence of China's Family Planning Policy. Deaths are determined by the life expectancy, which is modeled as a function of human well-being

(Equations S3-S6). Migration is calculated based on historical statistical data and per capita GDP differences between YRB and the national average level (Equations S7-S8).

2.2.2. Economy

L267-268: Capital elasticities are obtained from T21-China model. Given that they are time-varying, it would useful for readers to understand the assumptions in that model for deriving the shape. The lookup table stops at 2020. Is the parameter value then held constant for future projections? If so, is that a valid assumption for projecting future production?

Response： Thank you for the comments.

The capital elasticities in our model were originally based on the T21-China model but were modified and calibrated against historical data. The time-varying shape reflects China's economic structural transformation. For future projections (after 2020), the parameters were held constant at the 2020 the value. We consider this a valid and robust choice to avoid the high uncertainty associated with unpredictable changes in production technology over the long term.

We have revised the text to be more accurate:

α_s and $1\text{-}\alpha_s$ are capital and labor elasticities from T21-China (Qu et al., 2020) and calibrated using historical data;

L264: TFP is an important yet endogenous variable in the model, which should also be documented as an equation in the manuscript given that it includes several effects (infrastructure, education, health). The elasticities for these effects, unlike labour and capital, are constant parameters in the model. How were these elasticities determined? Education (avg years of schooling) is derived from census yearbooks for the historical period, but what is the strategy for future projections? Infrastructure is also an exogenous variable, however, the data and its source does not seem to be documented in the manuscript or the supplement.

Response： Thank you for the comments. Our response to each point is as follows:

TFP equations: Due to space limit in the main text (we can only list a few equations for each sector), the detailed mathematical formulation of TFP (including the effects of infrastructure, education, and health) is provided in the Supporting Information (Sector S2.2 Equations S10-S12). We have now added a clear citation in the manuscript Section 2.2 to explicitly direct readers to these equations. The revised texts are shown below:

TFPs is the total factor productivity calculated from exogenous (infrastructure, education, and elasticity) and endogenous variables (health and labor force) (Equations S10-S12);

Elasticities: The elasticities for infrastructure, education, and health were adopted directly from the T21-China model (Qu et al., 2020). Given the robust validation of the T21-China model in simulating national-scale development dynamics, we directly used these parameter values without localization given the lack of specific regional coefficients for these long-term structural factors. The explanation has been added in the revised Supporting Information as:

Owing to data limitations at the regional scale, both the infrastructure variables and the TFP elasticities are directly taken from the T21-China model (Qu et al., 2020).

Future projections for Education: For the projection of average years of schooling, there is an upper limit, so we employed a growth strategy with saturation. The saturation level (peak) was set based on the highest value observed globally in 2023 from Germany. The variable grows according to the annual growth rates specified in China's national planning documents until it reaches this peak, after which it is assumed to remain constant. This projection process is explained in the revised Supporting Information Sector S6 as:

The average years of schooling for the Yellow River Basin was projected based on targets outlined in China's Five-Year Plan. Since the average years of schooling has an upper limit, the maximum value was set to the highest reported in 2023 for Germany (14.3 years, data from https://ourworldindata.org/grapher/average-years-of-schooling). Specifically, we assumed that all nine provinces would meet the national target of 11.3 years by 2025 as outlined in the 14th Five-Year Plan. For years after 2025, the average years of schooling was set to increase at an annual rate of 0.114 years (equivalent to 0.57 years per five-year period) until the upper limit is reached.

Infrastructure data: The infrastructure density data was also sourced from the T21-China model. This data was not localized due to the lack of consistent, long-term historical time series for the YRB region. In the revised Supporting Information, we have now explicitly clarified this data source and the rationale for using the T21-China dataset as a proxy. The revised texts are shown below:

$$Inf_{eff} = \left(\frac{IFD}{IniIFD}\right)^{\epsilon_3} \tag{S13}$$

Inf_{eff} represents the effect of infrastructure on total factor productivity, IFD is the infrastructure density, with initial value from $IniIFD$; ϵ_1, ϵ_2 and ϵ_3 are the elasticities of TFP to education and health. Owing to data limitations at the regional scale, both the infrastructure variables and the TFP elasticities are directly taken from the T21-China model (Qu et al., 2020).

L270: Crop and livestock prices are exogenous inputs from China statistical yearbook up to 2020. Are prices then held constant for the future period?

Response: Thank you for the comments.

Yes, all prices in the model are held constant for the future projections. Given the inherent unpredictability and stochastic nature of agricultural market prices, we assumed they remain at the 2020 level. This is a standard assumption in long-term simulation models to avoid introducing arbitrary assumptions about future market trends.

Within the Economy sector of the model, there is internal feedback between GDP and the Capital stocks through private and government savings as a fraction of GDP. I recommend including this link in Figure 3 with a brief description in the manuscript. Currently, in the supplement, the formulations surrounding private and governmental investments are not fully described. There are several lookup tables here (e.g., investment share tables, saving share of income table), which

would benefit from some elaboration in the supplement at the very least.

Response: We apologize for the omission of these critical details in the previous version. We have corrected this by updating Figure 4 to visualize this feedback loop (see below):

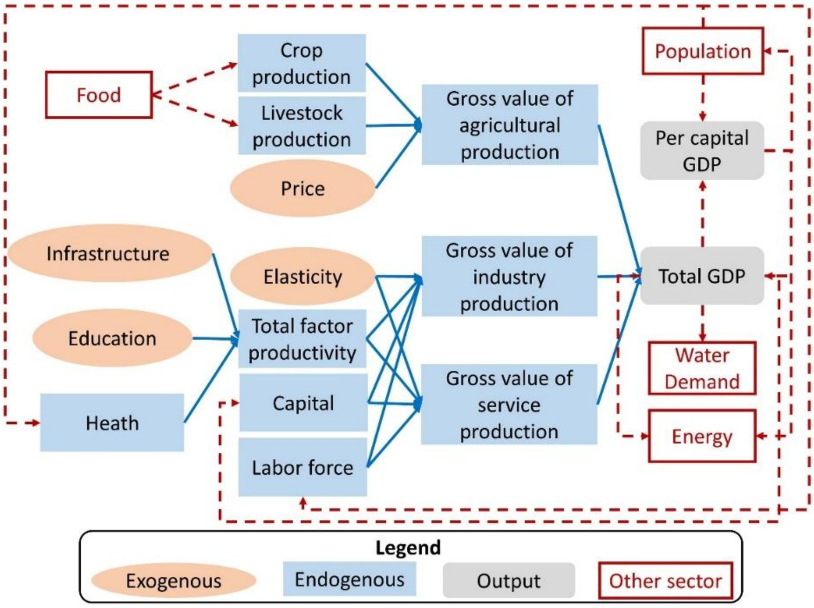

Fig. 1 (Fig. 4) Structure of the *Economy* sector. Blue lines indicate connections inside the sector, red dotted lines indicate connections with other sectors.

Furthermore, we have added the complete mathematical derivations for the investment significantly in the revised Supporting Information, as shown below:

The $Invest_s$ is calculated as

$$Invest_s = Invest_{total} \times InvestShare_s \tag{S14}$$

where $Invest_{total}$ is the total investment, $InvestShare_s$ is the investment share of each industry type.

Total investment ($Invest_{total}$) includes private investment ($Invest_{private}$) and government investment ($Invest_{govern}$):

$$Invest_{total} = Invest_{private} + Invest_{govern} \tag{S15}$$

$$Invest_{private} = \sum_{s=1}^{3} GDP_s \times Share_{DisInc} \times Share_{saving} \tag{S16}$$

$$Invest_{govern} = (\sum_{s=1}^{3} GDP_s) \times Share_{GovExp} \times Share_{WithoutInt} \times Share_{GovInv} \tag{S17}$$

where $Share_{DisInc}$ is the ratio of disposable income to the total GDP ($\sum_{s=1}^{3} GDP_s$), from the China Statistical Yearbook (NBSC, 2020c), $Share_{saving}$ is the ratio of savings to disposable income, from T21-China model (Qu et al., 2020); $Share_{GovExp}$ is the ratio of government expenditure to the total GDP, from the China Statistical Yearbook (NBSC, 2020c), $Share_{WithoutInt}$ is the ratio of government expenditure without interest to the government expenditure, $Share_{GovInv}$ is the ratio of government investment to the government expenditure without interest, both from T21-China model (Qu et al., 2020).

2.2.3. Energy

Most of the endogenous variables in the energy sector are formulated as linear functions of GDP per capita. Though this was simply stated as such in the manuscript and supplement, there is no justification as to why linear relationships were chosen. More concerning, is the use of IF THEN ELSE functions to specify two different sets of constants and slopes before and after 2010 within a few of these relationships. These assumptions are not documented either.

Response: Thank you for the comments.

The choice of linear functions was driven by empirical analysis of historical data. Our regression results demonstrated that linear models provided the best fit for the relationship between energy consumption and economic indicators. Residential electricity consumption showed a strong linear correlation with GDP per capita. Fossil fuel consumption (coal, oil, gas) was linearly correlated with sectoral GDP.

The historical data showed a distinct structural change around 2010, reflecting a shift in China's socio-economic development and energy policies (e.g., improved energy efficiency and industrial restructuring). Consequently, the slope of these linear relationships changed significantly after 2010. To accurately capture this regime shift, we employed a piecewise linear regression approach (implemented via IF THEN ELSE logic), fitting separate coefficients for the pre-2010 and post-2010 periods. This approach significantly improved the model's historical accuracy compared to a single continuous function.

We have clarified these assumptions in the revised Supporting Information Section S2.3. The revised texts are shown below:

$Para_{IC}$ and $Const_{IC}$ represent the slope and intercept derived from the linear fit of historical industry GDP and coal consumption. Similarly, $Para_{ISO}$, $Const_{ISO}$ and $Para_{ISG}$, $Const_{ISG}$ denote the corresponding parameters for oil and gas, respectively. Notably, these parameters were estimated using a piecewise linear regression with distinct coefficients for the pre- and post-2010 periods to capture the structural changes around 2010 driven by socio-economic shifts.

2.2.4. Food

L293-297: Per capita meat demand is derived from a nonlinear regression of GDP per capita and historical per capita production. In the model, there is an IF THEN ELSE function which specifies a different relationship for time 1981. Why is that so?

Response: Thank you for the comments.

The year 1981 serves as the initial simulation year of the model. The IF THEN ELSE function is to provide the initial value for per capita meat demand. This choice is purely a technical reason under VENSIM.

In the model, a circular dependency exists within a single time step: per capita GDP influences per capita meat demand, and meat demand determines meat production (assuming supply equals demand historically), agriculture output, and therefore total GDP, which, in turn, calculates per capita GDP.

Since these variables influence each other simultaneously, VENSIM identifies this as a simultaneous equation loop. Without a defined initial state, the software cannot resolve the loop for the first time step (1981) and triggers a computation error. Therefore, the function explicitly assigns an initial value to per capita meat demand at Time = 1981 based on historical data. This breaks the circular loop at the start of the simulation, allowing the endogenous feedback calculations to proceed correctly for all subsequent years.

L303-305: Crop yields are influenced by precipitation, temperature and CO2 concentration from the Climate sector with a reference to the supplement. The supplement, however, does not describe the influence of CO2 concentrations. There is also a mismatch between the model and the supplement creating confusions. In the model, there are two different variables: 'YRB food production' and 'province crop production' determined by 'province yield' and 'province yield future' respectively. Precipitation and temperature influences only 'province yield' whereas CO2 concentration influences 'province yield future'. The supplement does not sufficiently describe this formulations for me to understand the conceptualization here. Please clarify.

Response: We apologize for the confusion caused by the variable naming inconsistency and the incomplete description in the Supporting Information.

Province Yield represents historical yield in each province, derived from historical regression models, driven by fertilizer use, temperature, and precipitation. It serves as the baseline yield for the historical period (1981–2020) and the foundation for future projections. Province Yield Future represents the final yield for the future (2021–2100) after incorporating the CO_2 fertilization effect.

Historical yield trends are driven by climate and management factors while CO2 played a relatively minor role compared to the dominant technological such as fertilizer. Therefore, the specific contribution of CO_2 was not isolated in historical yield simulation. However, for future scenarios, atmospheric CO_2 concentrations diverge significantly from historical trends and become more important factor for yield. To capture this, we introduce an explicit CO_2 fertilization factor for future crop yield projection.

We have updated the Supporting Information Section S2.4 to explicitly describe this formulation. The revised texts are shown below:

Since the future crop yield ($YieldFut_c$) is increasingly affected by CO_2 fertilization, the CO_2 fertilization effect is explicitly modeled with distinction between C3 and C4 crops. The $YieldFut_c$ is calculated as:

$$YieldFut_c = \begin{cases} 1 + (0.000964 \times (CO_2 - 413.2) \times Yield_c, c = other\ \ crops \\ 1 + (0.0000545 \times (CO_2 - 413.2) \times Yield_c,\ c = maize \end{cases}$$

(S35)

where 0.000964%/ppm is the CO_2 fertilization coefficient for C3 crops (i.e., rice, wheat, beans, potato, cotton and oil plants) (Kimball, 2016); 0.0000545 %/ppm is the CO_2 fertilization coefficient for the C4 crop maize (Rezaei et al., 2023); CO_2 is the atmospheric CO_2 concentration, and 413.2 ppm is the CO_2 concentration in 2020.

Food demand formulations (e.g., S34 and S36) rely on exogenous inputs for daily requirements and dietary proportions from the historical period. How is future demand projected beyond the historical period?

Response： Similar to other exogenous parameters, the exogenous inputs (daily requirements and dietary proportions) are held constant at the most recent historical values. This assumes that without specific external interventions, the dietary structure maintains its current pattern. Importantly, these variables are designed to be adjustable within the model, allowing us to modify them to create various alternative dietary scenarios. For example, we can adjust the dietary proportions to design a diet shift scenario towards a "balanced diet" or "reduced meat consumption", enabling us to analyze the impacts of different dietary choices.

2.2.8. Land

The transfer matrix stock is used to calculate the 'final historical transfer matrix' in the model. There is a separate 'future transfer matrix' variable with multiple subscripts having the value 0 while others simply take the final historical transfer matrix value. This warrants elaboration in the supplement, since it's not clear how land use changes are determined beyond data replication for the historical period.

Response: We apologize for the omission of these critical details in the previous version. The "Future Transfer Matrix" is indeed distinct from the historical one because it incorporates specific transition rules and constraints aligned with the policies assumed in the future baseline scenario. The "zero values" in the matrix impose specific restrictions on land-use conversion. Due to strict ecological protection, we assume these three ecological land types with dense vegetation cover (cropland, forest, and grassland) can convert among themselves, but are restricted from converting into other types (hence, transfer probabilities to other categories are set to 0). The unused land is allowed to convert into cropland, forest, grassland, or settlement, but no lands are allowed to convert back into unused land. The wetland area is held constant at 2020 levels (no conversion in or out).

We have added the future transition assumptions to Table 2 in the main text and the Supporting Information Section S6:

Outline of the National Overall Planning on Land Use (2006 - 2020) (https://www.ndrc.gov.cn/fggz/fzzlgh/gjjzxgh/201705/t20170517_1196768_ext.html) requires a minimum cropland area for each province to prevent excessive farmland loss. The minimum threshold, denoted as the red line for cropland area, is included in the model's future projection. For other land uses, ecological lands with dense vegetation cover (cropland, forest, grassland) can convert into each other but not into the rest land due to ecological policies; unused land can be converted into the rest land, but not the other way around; and wetland area remains unchanged.

3. Model validation and application

I appreciate that the authors have provided the scenario specification in the supplement S5 along with Table 2 that lists some of the data sources and assumptions for future baseline of key variables. I understand that there are over a 100 exogenous variables in the model. However, it is quite disorienting that there is an excel file for time-varying data inputs and several other time-varying inputs hidden within lookup tables (i.e., data points that are interpolated). And, as alluded to in my above comments in the model description, it is unclear how many of these time-varying inputs are formulated for the future baseline (most stop at historical period). I strongly recommend that an additional supplement file is included for ALL exogenous time-varying inputs, including those hidden in table

functions. For clarity, the file should list all the historical data points used to drive model behaviour along with a column for data source, followed by all the future baseline data points with another column for data source/assumptions.

Response: Thank you for this suggestion regarding the model data files.

We have restructured the exogenous time-varying data inputs into a single file named 'exogenous_time-varying_inputs.xlsx', which comprises three sheets. The first sheet, 'NOTES', lists the references; the second sheet, 'history', contains historical input values and their sources; and the third sheet, 'future', details future input values along with their sources and corresponding assumptions. The file is available at https://doi.org/10.5281/zenodo.17568963 or https://github.com/sangshan-ss/CHANS_SD_YRB/blob/main/exogenous_data_info/exogenous_time-varying_inputs.xlsx.

Sector	Variables	Unit	Sources/Assumptions
population	fertility[Shanxi]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	fertility[Neimeng]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	fertility[Shandong]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	fertility[Henan]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	fertility[Sichuan]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	fertility[Shaanxi]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	fertility[Gansu]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	fertility[Qinghai]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	fertility[Ningxia]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	proportion of male babies[Shanxi]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	proportion of male babies[Neimeng]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	proportion of male babies[Shandong]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	proportion of male babies[Henan]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	proportion of male babies[Sichuan]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	proportion of male babies[Shaanxi]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	proportion of male babies[Gansu]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	proportion of male babies[Qinghai]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	proportion of male babies[Ningxia]	dmnl	United Nations World Population Prospects 2024 (https://population.un.org/wpp/)
population	urban pop share province[Shanxi]	dmnl	Upper limit of 80% from the Report on the Analysis and Forecast of China's Urbanization Trend toward Modernization (2020) and growth rates (1%/yr) specified in China's Five-Year Plans.
population	urban pop share province[Neimeng]	dmnl	Upper limit of 80% from the Report on the Analysis and Forecast of China's Urbanization Trend toward Modernization (2020) and growth rates (1%/yr) specified in China's Five-Year Plans.
population	urban pop share province[Shandong]	dmnl	Upper limit of 80% from the Report on the Analysis and Forecast of China's Urbanization Trend toward Modernization (2020) and growth rates (1%/yr) specified in China's Five-Year Plans.
population	urban pop share province[Henan]	dmnl	Upper limit of 80% from the Report on the Analysis and Forecast of China's Urbanization Trend toward Modernization (2020) and growth rates (1%/yr) specified in China's Five-Year Plans.
population	urban pop share province[Sichuan]	dmnl	Upper limit of 80% from the Report on the Analysis and Forecast of China's Urbanization Trend toward Modernization (2020) and growth rates (1%/yr) specified in China's Five-Year Plans.
population	urban pop share province[Shaanxi]	dmnl	Upper limit of 80% from the Report on the Analysis and Forecast of China's Urbanization Trend toward Modernization (2020) and growth rates (1%/yr) specified in China's Five-Year Plans.
population	urban pop share province[Gansu]	dmnl	Upper limit of 80% from the Report on the Analysis and Forecast of China's Urbanization Trend toward Modernization (2020) and growth rates (1%/yr) specified in China's Five-Year Plans.
population	urban pop share province[Qinghai]	dmnl	Upper limit of 80% from the Report on the Analysis and Forecast of China's Urbanization Trend toward Modernization (2020) and growth rates (1%/yr) specified in China's Five-Year Plans.
population	urban pop share province[Ningxia]	years	Upper limit of 80% from the Report on the Analysis and Forecast of China's Urbanization Trend toward Modernization (2020) and growth rates (1%/yr) specified in China's Five-Year Plans.
population	average years of schooling[Shanxi]	years	The maximum value reported for Germany in 2023 (14.3 years), increase at an annual rate of 0.114 years (equivalent to 0.57 years per five-year period) until the upper limit is reached.
population	average years of schooling[Neimeng]	years	The maximum value reported for Germany in 2023 (14.3 years), increase at an annual rate of 0.114 years (equivalent to 0.57 years per five-year period) until the upper limit is reached.
population	average years of schooling[Shandong]	years	The maximum value reported for Germany in 2023 (14.3 years), increase at an annual rate of 0.114 years (equivalent to 0.57 years per five-year period) until the upper limit is reached.
population	average years of schooling[Henan]	years	The maximum value reported for Germany in 2023 (14.3 years), increase at an annual rate of 0.114 years (equivalent to 0.57 years per five-year period) until the upper limit is reached.
population	average years of schooling[Sichuan]	years	The maximum value reported for Germany in 2023 (14.3 years), increase at an annual rate of 0.114 years (equivalent to 0.57 years per five-year period) until the upper limit is reached.
population	average years of schooling[Shaanxi]	years	The maximum value reported for Germany in 2023 (14.3 years), increase at an annual rate of 0.114 years (equivalent to 0.57 years per five-year period) until the upper limit is reached.
population	average years of schooling[Gansu]	years	The maximum value reported for Germany in 2023 (14.3 years), increase at an annual rate of 0.114 years (equivalent to 0.57 years per five-year period) until the upper limit is reached.
population	average years of schooling[Qinghai]	years	The maximum value reported for Germany in 2023 (14.3 years), increase at an annual rate of 0.114 years (equivalent to 0.57 years per five-year period) until the upper limit is reached.
population	average years of schooling[Ningxia]	years	The maximum value reported for Germany in 2023 (14.3 years), increase at an annual rate of 0.114 years (equivalent to 0.57 years per five-year period) until the upper limit is reached.
population	total agriculture employment share[Shanxi]	dmnl	Keep the historical values for the last year unchanged.
population	total agriculture employment share[Neimeng]	dmnl	Keep the historical values for the last year unchanged.
population	total agriculture employment share[Shandong]	dmnl	Keep the historical values for the last year unchanged.
population	total agriculture employment share[Henan]	dmnl	Keep the historical values for the last year unchanged.
population	total agriculture employment share[Sichuan]	dmnl	Keep the historical values for the last year unchanged.
population	total agriculture employment share[Shaanxi]	dmnl	Keep the historical values for the last year unchanged.
population	total agriculture employment share[Gansu]	dmnl	Keep the historical values for the last year unchanged.
population	total agriculture employment share[Qinghai]	dmnl	Keep the historical values for the last year unchanged.

Fig2 Screenshot of the excel

Table S2 lists quite a number of exogenous inputs. I would like for a distinction between exogenous variables (time-varying) and parameters (constants). While the above file should contain all the historical and future data points for exogenous variables, another file or table should be provided for all model parameters along with the estimated value and data source (literature or model calibration). Having these files would considerably improve the transparency of the model – especially

for readers unfamiliar with Vensim software.

Response： Thank you for this suggestion regarding the model data files.

According to the suggestion, we have compiled all time-invariant exogenous variables into a spreadsheet file named 'exogenous_parameters_(constant)_inputs.xlsx'. This file consists of 10 sheets. The first sheet, labeled 'NOTES', contains two explanatory notes and the full list of references. The remaining nine sheets are named after their respective sectors and list the constant parameters and data sources for each sector. The file is available at https://doi.org/10.5281/zenodo.17568963 or https://github.com/sangshan-ss/CHANS_SD_YRB/blob/main/exogenous_data_info/exogenous_parameters_(constant)_inputs.xlsx.

	A	B	C	D	E	F
1	**Variables (Unit)**	**Sources**	**Value**			
2	inflection point (dmnl)	Fitted by historical data	-1.56528			
3	steepness (dmnl)	Fitted by historical data	6.34527			
4	min life expectancy (years old)	Fitted by historical data	56.1739			
5	max life expectancy (years old)	Fitted by historical data	80			
6	reference saturation PC GDP (CNY)	T21-China (Qu et al., 2020)	254171			
7	reference saturation years of schooling (years)	T21-China (Qu et al., 2020)	16			
8	PC GDP weight (dmnl)	T21-China (Qu et al., 2020)	0.7			
9	**age**		**15**	**16**	**17**	**18**
10	age specific fertility distribution 1986 (dmnl)	China Population Census Yearbook	0.000128586	0.000892	0.003654	0.011009
11	age specific fertility distribution 2010[shanxi] (dmnl)	China Population Census Yearbook	0	0	0.00032	0.00052
12	age specific fertility distribution 2010[neimeng] (dmnl)	China Population Census Yearbook	0	0.00151	0.00234	0.00455
13	age specific fertility distribution 2010[shandong] (dmnl)	China Population Census Yearbook	0	0.00025	0.00149	0.00292
14	age specific fertility distribution 2010[henan] (dmnl)	China Population Census Yearbook	0	0.00025	0.00149	0.00292
15	age specific fertility distribution 2010[sichuan] (dmnl)	China Population Census Yearbook	0	0.00151	0.00234	0.00455
16	age specific fertility distribution 2010[shaanxi] (dmnl)	China Population Census Yearbook	0	0	0.00032	0.00052
17	age specific fertility distribution 2010[gansu] (dmnl)	China Population Census Yearbook	0	0.00151	0.00234	0.00455
18	age specific fertility distribution 2010[qinghai] (dmnl)	China Population Census Yearbook	0	0.00151	0.00234	0.00455
19	age specific fertility distribution 2010[ningxia] (dmnl)	China Population Census Yearbook	0	0.00151	0.00234	0.00455
20	age specific fertility distribution 2050 (dmnl)	T21-China (Qu et al., 2020)	0	0.00086	0.003	0.006
21	**age**		**0**	**1**	**2**	**3**
22	migration rate age proportion[male] (dmnl)	China Population Census Yearbook	0.00258541	0.006206	0.007012	0.008057
23	migration rate age proportion[female] (dmnl)	China Population Census Yearbook	0.00246273	0.006211	0.006914	0.007709
24	initial population[male,shanxi] (person)	China Population Census Yearbook	274450	229366	240709	257767
25	initial population[male,neimeng] (person)	China Population Census Yearbook	208297	174080	182689	195636
26	initial population[male,shandong] (person)	China Population Census Yearbook	785041	656081	688526	737320
27	initial population[male,henan] (person)	China Population Census Yearbook	788809	659230	691831	740859
28	initial population[male,sichuan] (person)	China Population Census Yearbook	776876	649257	681365	729652
29	initial population[male,shaanxi] (person)	China Population Census Yearbook	311085	259983	272840	292175
30	initial population[male,gansu] (person)	China Population Census Yearbook	210182	175655	184341	197405
31	initial population[male,qinghai] (person)	China Population Census Yearbook	41241	34466	36171	38734
32	initial population[male,ningxia] (person)	China Population Census Yearbook	41450	34641	36354	38930
33	initial population[female,shanxi] (person)	China Population Census Yearbook	245719	204983	216070	232771
34	initial population[female,neimeng] (person)	China Population Census Yearbook	186238	155362	163766	176424
35	initial population[female,shandong] (person)	China Population Census Yearbook	747619	623674	657409	708224
36	initial population[female,henan] (person)	China Population Census Yearbook	744337	620936	654523	705115
37	initial population[female,sichuan] (person)	China Population Census Yearbook	718904	599719	632159	681022
38	initial population[female,shaanxi] (person)	China Population Census Yearbook	282844	235952	248715	267940
39	initial population[female,gansu] (person)	China Population Census Yearbook	192186	160324	168996	182059
40	initial population[female,qinghai] (person)	China Population Census Yearbook	37945	31654	33366	35946
41	initial population[female,ningxia] (person)	China Population Census Yearbook	37945	31654	33366	35946
42						
43						
44						
45						

NOTES | population | economy | energy | food | water demand | water supply | sediment | land | carbon

Fig3 Screenshot of the excel

The main text for description of simulation results should explicitly refer to figures to guide the reader.

Response: Thank you for the suggestion.

In the revised manuscript, we have referred to all figures and tables.

The authors have emphasized high correlations for historical fits. However I am slightly skeptical since the current approach risks overfitting: several relationships are fitted to historical to the historical period being validated and many exogenous time series drive the dynamics of the model. To build confidence in the behavioural validity, the authors should use out-of-sample validation: model fit to a subset of historical time series (e.g. 1981 - 2005) and use the remains sample points to validate the model outputs.

Response: Thank you for pointing out this issue.

We agree with the reviewer that out-of-sample validation is a widely used method to evaluate model's performance. However, applying this method (e.g., calibrating on 1981–2005 and testing on 2006–2020) presents a fundamental challenge for the coupled human and natural systems in the Yellow River Basin due to significant structural changes and regime shifts driven by human intervention and policies (e.g., water use allocation, green-for-grain, ecological restoration projects, etc).

From the modeling perspective, the high correlations in historical validation are unlikely to be a result of overfitting. When modeling the human or natural processes, most of the qualitative relationships are constructed based on established theoretical frameworks (e.g., Cobb-Douglas production functions, physical hydrological laws) rather than data-mining techniques or high-order polynomial fitting prone to memorize noise. The parameters derived from simple linear or low-order non-linear structures have clear physical/economic meanings.

The high R values reflect the inherent characteristics of the CHANS systems. The slowly involving socio-economic variables (e.g., Population, GDP) have a strong inertia and show smooth development trends under a stable policy environment. These systems rarely exhibit sudden jumps, making them naturally easier to capture with high (R > 0.95 is common and expected in such sectors). In contrast, natural subsystems (e.g., runoff, sediment load) exhibit significantly lower R values. This is because they are driven by stochastic climatic disturbances (precipitation/temperature) and less predictable. If we were overfitting, all variables (including natural ones) would likely show artificially high fits. The discrepancy between the higher R in stable human systems and lower R in natural systems is consistent with our expectation due to distinct behaviors of these components.

Sensitivity analyses to quantify parametric uncertainty is also a validation standard in system dynamics. No such analysis appears in the manuscript or supplement.

Response: Thank you for pointing out this issue.

We have added the sensitivity analyses to the Supporting Information Sector S3. The revised texts are shown below:

Model sensitivity analyses were performed for all constant parameters across all sectors using the Multivariate Monte Carlo tool in Vensim DSS. There are > 500 constant parameters listed in the supplementary spreadsheet file "sensitivity analysis parameters.xlsx" available at https://doi.org/10.5281/zenodo.17568963 or https://github.com/sangshan-ss/CHANS_SD_YRB/blob/main/sensitivity_analysis_files/sensitivity_analysis_parameters.xlsx. We assigned a ±10% variation to each parameter, following a uniform distribution within this range. A total of 500 simulations with randomly generated parameter values were executed to assess how parametric uncertainty propagates to affect the model's future projections (Fig. S10).

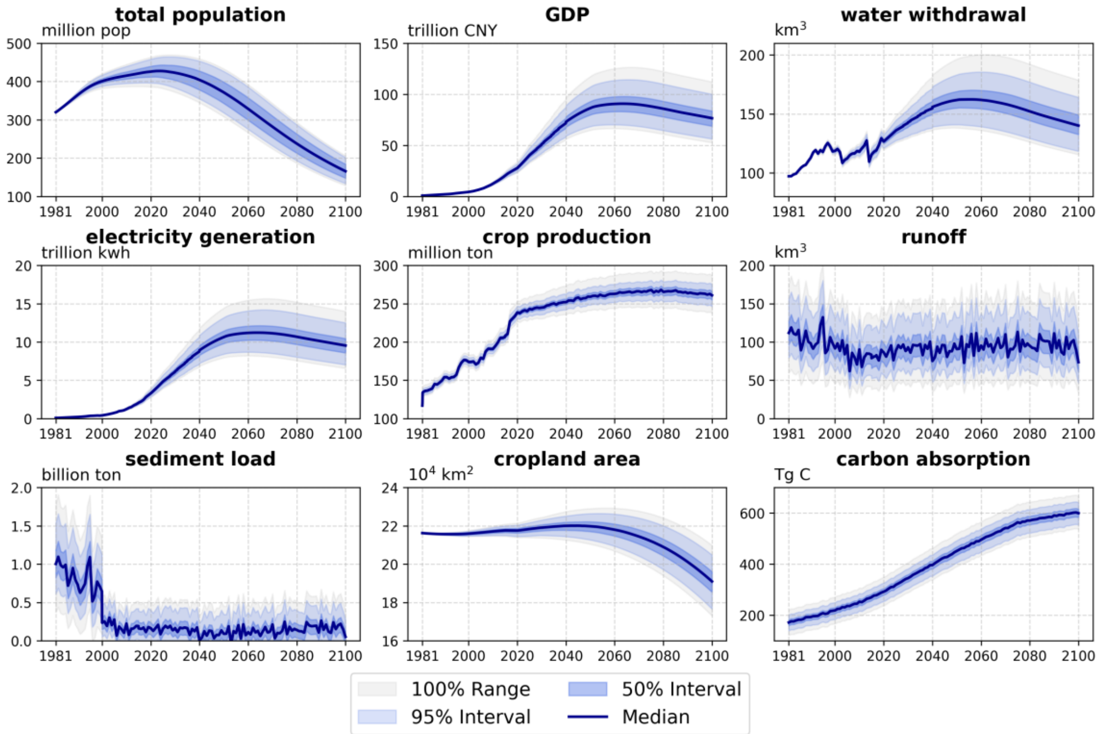

Fig 4 (Fig S10) Sensitivity analyses of the model's key indicators to model parameters in the future baseline projections (1981–2100). The solid dark blue line represents the median of 500 simulations. The shaded areas, from dark to light, correspond to the 50% interval, 95% interval, and the full 100% range of outcomes, respectively.

Key indicators of human sectors, such as GDP and electricity generation, exhibited a characteristic 'fan-shaped' divergence, with the uncertainty range widening significantly after 2040. This indicates that these variables are highly sensitive to the socio-economic parameters through accumulation effects. Despite this divergence, the median trajectories showed robust growth followed by stabilization or gradual decline, confirming that the economic transitions are structurally plausible. In contrast, carbon absorption and agricultural production exhibited relatively narrower confidence intervals, demonstrating a high structural stability. The confidence intervals of hydro-climatic variables, particularly runoff and sediment load, reflected high-frequency interannual fluctuations driven by climatic inputs rather than structural instability. Overall, despite growing uncertainty in the long-term projections, the simulation envelopes (100% range) remain within realistic bounds, with no catastrophic collapse (e.g., negative values), thereby confirming the model's behavioral robustness.

A summary of the sensitivity analysis results has been added to the manuscript Section 3.1. The revised texts are shown below.

Sensitivity analyses with randomly generated parameter values reveal growing uncertainty in the long-term trajectories of socio-economic and natural variables; the model maintains behavioral robustness across all runs without catastrophic collapse or unrealistic oscillations (see Supporting Information S3 for details).

L436: What are some plausible explanations for Shanxi exhibiting lower accuracy?

Response: During the revision, we re-evaluated the Capital Elasticity and adjusted it to > 0.3 (a more scientifically justifiable range) to enhance the model's physical realism. We also corrected an error in the retirement age parameter. These adjustments resulted into a decrease R in validation metrics for Neimeng, making Neimeng exhibiting lower simulation accuracy.

As a resource-dependent region dominated by the coal industry, Neimeng's historical economic data exhibit significant fluctuations driven by energy market cycles. The model parameters derived from long-term structural trends inevitably smooths out these high-frequency market shocks. In addition, Neimeng has undergone aggressive industrial restructuring and ecological restoration (e.g., Grain for Green) in the basin. These discrete policy shocks are challenging to reproduce perfectly.

We have added these explanations in the revised manuscript. The revised texts are shown below:

Some provinces, like Neimeng, exhibit lower simulation accuracy for certain indicators than others, probably due to the simplification of relevant modelled processes, imperfect parameterizations, and external policy interventions.

L504: What is the red line? I do not see a red line in Figure 14(a) land area. Why does cropland area "need" to stay above the red line?

Response: Thank you for pointing out this issue.

The "Cropland Red Line" is a specific policy term in China, referring to the mandatory minimum threshold for cropland area set by the central government for each province to guarantee national food security. The cropland area has to stay above this line because it is a strict legal and administrative constraint; falling below this limit is strictly prohibited by land management regulations.

We apologize for the confusion about the absence in Figure 14(a). The red line was not visually plotted in Figure 14(a) because the figure is a stacked line chart displaying the cumulative composition of six different land-use types. Overlaying a threshold line for a single category (cropland) on a stacked chart would be visually misleading and difficult to interpret. Although not shown in the figure, this constraint is embedded in the model structure. We set the "Red Line" as a hard boundary condition (lower limit) in the land-use module. Therefore, the simulation logic inherently prevents the cropland area from dropping below this policy threshold in all future scenarios.

We have added the explanation in the revised manuscript. The revised texts are shown below:

Under the future baseline scenario, land use patterns remain relatively stable based on historical trends, as no new land policies are introduced and cropland area has to stay above the red line (a mandatory minimum cropland area for each province for food security) (Fig.14 (a)).

In the Supporting Information, we also added the explanation. The revised texts are shown below:

Outline of the National Overall Planning on Land Use (2006 - 2020) (https://www.ndrc.gov.cn/fggz/fzzlgh/gjjzxgh/201705/t20170517_1196768_ext.html) requires a minimum cropland area for each province to prevent excessive farmland loss. The minimum threshold, denoted as the red line for cropland area, is included in the model's future projection. For other land uses, ecological lands with dense vegetation cover (cropland, forest, grassland) can convert into each other but not into the rest land due to ecological policies; unused land can be converted into the rest land, but not the other way around; and wetland area remains unchanged.

4. Discussion

L635-638: Although the model is presented as a CHANS framework, the Climate sector functions as an exogenous driver rather than a dynamically coupled component. The absence of climate feedbacks limits the scope of coupling. The model cannot represent second-order feedback dynamics such as how afforestation or carbon emissions might alter local hydrology or climate, critical for CHANS theory. Therefore, the authors should explicitly acknowledge this limitation and avoid overstating its scope, especially in L597-597 where they state the "model's comprehensive coupling broadens its scope of application." They should also discuss implications of climate feedback omissions for scenario analysis.

Response： We agree that the absence of climate feedbacks and other possible feedback is a limitation. Accurately modeling these climate feedbacks at the basin scale requires using specialized climate models, which is technically infeasible to incorporate them into current model structure. We have modified the statement to be narrow the model's scope:

Our model's integration of socio-economic and ecological sectors provides a robust framework for addressing regional CHANS challenges and offers practical guidance for sustainable development.

We also have added a new paragraph in the Discussion section to explicitly discuss the implications of omitting climate feedbacks:

Moreover, there are still important feedbacks absent from the current coupling framework. Notable examples include the effects of land use change on climate, the effects of climate change on economic growth, and the effects of pricing on energy use and carbon emissions. These missing feedbacks could be incorporated based on recent studies, including the land use feedback on precipitation through moisture recycling (Sang et al., 2025), socioeconomic losses from climate change impacts (Waidelich et al., 2024b), and integration with Computable General Equilibrium (CGE) models

(Fujimori et al., 2014).

The authors have justified using CMIP6 data instead of observed records to ensure temporal consistency, but they do not discuss the potential bias introduced by replacing observed data. In the model, it appears that the authors have considered SSP1.26, 2.45 and 5.85. Showing and discussing the sensitivity of results to alternative climate pathways could help demonstrate robustness.

Response： We acknowledge the concern regarding influence of CMIP6 climate data. First, we performed bias correction on the CMIP6 raw data based on climate observations from 1981 to 2014. This process aligns the statistical properties (mean) of the model data with the observed records, effectively removing the systemic bias of the climate model while preserving the temporal consistency required for long-term simulation. We apologize for omitting the description of the bias correction process in the previous version of the manuscript. We have revised the manuscript Section 3.2.1 and Supporting Information Section S6. The revised texts are show below:

To reduce systemic biases in the raw CMIP6 data, we applied bias correction using CN05 and ground weather stations observations from 1981 to 2014. By statistically aligning the mean of the CMIP6 data with observation records, systemic bias is removed while retaining the temporal consistency required for long-term simulation.

Technically, the model is capable of simulating climate data from SSP1-2.6, SSP2-4.5, and SSP5-8.5 scenarios. For the future baseline scenario presented, we selected SSP2-4.5 because it represents a "middle-of-the-road" trajectory that align with the ongoing real-world climate change. Exploring the sensitivity to other pathways (SSP1-2.6 and SSP5-8.5) is a valuable direction for future scenario comparisons.

References:

Breach, P. A. and Simonovic, S. P.: ANEMI3: An updated tool for global change analysis, PLoS ONE, 16, e0251489, https://doi.org/10.1371/journal.pone.0251489, 2021.

Calvin, K. and Bond-Lamberty, B.: Integrated human-earth system modeling—state of the science and future directions, Environ. Res. Lett., 13, 063006, https://doi.org/10.1088/1748-9326/aac642, 2018.

Cheng, Y., Huang, M., Lawrence, D. M., Calvin, K., Lombardozzi, D. L., Sinha, E., Pan, M., and He, X.: Future bioenergy expansion could alter carbon sequestration potential and exacerbate water stress in the United States, Science Advances, 8, eabm8237, https://doi.org/10.1126/sciadv.abm8237, 2022.

Di Vittorio, A. V., Sinha, E., Hao, D., Singh, B., Calvin, K. V., Shippert, T., Patel, P., and Bond-Lamberty, B.: E3SM-GCAM: A Synchronously Coupled Human Component in the E3SM Earth System Model Enables Novel Human-Earth Feedback Research, Journal of Advances in Modeling Earth Systems, 17, e2024MS004806, https://doi.org/10.1029/2024MS004806, 2025.

Dobbs, G. R., Liu, N., Caldwell, P. V., Miniat, C. F., Sun, G., Duan, K., and Bolstad, P. V.: Inter-basin surface water transfers database for public water supplies in conterminous United States, 1986–2015, Sci Data, 10, 255, https://doi.org/10.1038/s41597-023-02148-5, 2023.

Dottori, F., Szewczyk, W., Ciscar, J.-C., Zhao, F., Alfieri, L., Hirabayashi, Y., Bianchi, A., Mongelli, I., Frieler, K., Betts, R. A., and Feyen, L.: Increased human and economic losses from river flooding with anthropogenic warming, Nature Clim Change, 8, 781–786, https://doi.org/10.1038/s41558-018-0257-z, 2018.

Forrester, J. W.: Principles of systems, Pegasus Communications, Inc, Waltham, MA, 1968.

Fu, B. and Li, Y.: Bidirectional coupling between the Earth and human systems is essential for modeling sustainability, National Science Review, 3, 397–398, https://doi.org/10.1093/nsr/nww094, 2016.

Fujimori, S., Masui, T., and Matsuoka, Y.: Development of a global computable general equilibrium model coupled with detailed energy end-use technology, Applied Energy, 128, 296–306, https://doi.org/10.1016/j.apenergy.2014.04.074, 2014.

Fujimori, S., Wu, W., Doelman, J., Frank, S., Hristov, J., Kyle, P., Sands, R., van Zeist, W.-J., Havlik, P., Domínguez, I. P., Sahoo, A., Stehfest, E., Tabeau, A., Valin, H., van Meijl, H., Hasegawa, T., and Takahashi, K.: Land-based climate change mitigation measures can affect agricultural markets and food security, Nat Food, 3, 110–121, https://doi.org/10.1038/s43016-022-00464-4, 2022.

Hasegawa, T., Sakurai, G., Fujimori, S., Takahashi, K., Hijioka, Y., and Masui, T.: Extreme climate events increase risk of global food insecurity and adaptation needs, Nat Food, 2, 587–595, https://doi.org/10.1038/s43016-021-00335-4, 2021.

Jain, S., Mindlin, J., Koren, G., Gulizia, C., Steadman, C., Langendijk, G. S., Osman, M., Abid, M. A., Rao, Y., and Rabanal, V.: Are We at Risk of Losing the Current Generation of Climate Researchers to Data Science?, AGU Advances, 3, e2022AV000676, https://doi.org/10.1029/2022AV000676, 2022.

Kimball, B. A.: Crop responses to elevated CO2 and interactions with H2O, N, and temperature, Current Opinion in Plant Biology, 31, 36–43, https://doi.org/10.1016/j.pbi.2016.03.006, 2016.

Li, X., Cheng, G., Lin, H., Cai, X., Fang, M., Ge, Y., Hu, X., Chen, M., and Li, W.: Watershed System Model: The Essentials to Model Complex Human-Nature System at the River Basin Scale, Journal of Geophysical Research: Atmospheres, 123, 3019–3034, https://doi.org/10.1002/2017JD028154, 2018.

Liu, J.: Integration across a metacoupled world, E&S, 22, art29, https://doi.org/10.5751/ES-09830-220429, 2017.

Liu, J., Dietz, T., Carpenter, S. R., Alberti, M., Folke, C., Moran, E., Pell, A. N., Deadman, P., Kratz, T., Lubchenco, J., Ostrom, E., Ouyang, Z., Provencher, W., Redman, C. L., Schneider, S. H., and Taylor, W. W.: Complexity of Coupled Human and Natural Systems, Science, 317, 1513–1516, https://doi.org/10.1126/science.1144004, 2007.

Monier, E., Paltsev, S., Sokolov, A., Chen, Y.-H. H., Gao, X., Ejaz, Q., Couzo, E., Schlosser, C. A., Dutkiewicz, S., Fant, C., Scott, J., Kicklighter, D., Morris, J., Jacoby, H., Prinn, R., and Haigh, M.: Toward a consistent modeling framework to assess multi-sectoral climate impacts, Nat Commun, 9, 1–8, https://doi.org/10.1038/s41467-018-02984-9, 2018.

Motesharrei, S., Rivas, J., Kalnay, E., Asrar, G. R., Busalacchi, A. J., Cahalan, R. F., Cane, M. A., Colwell, R. R., Feng, K., Franklin, R. S., Hubacek, K., Miralles-Wilhelm, F., Miyoshi, T., Ruth, M., Sagdeev, R., Shirmohammadi, A., Shukla, J., Srebric, J., Yakovenko, V. M., and Zeng, N.: Modeling sustainability: population, inequality, consumption, and bidirectional coupling of the Earth and Human Systems, National Science Review, 3, 470–494, https://doi.org/10.1093/nsr/nww081, 2016.

Qu, W., Shi, W., Zhang, J., and Liu, T.: T21 China 2050: A Tool for National Sustainable Development Planning, Geography and Sustainability, 1, 33–46, https://doi.org/10.1016/j.geosus.2020.03.004, 2020.

Rajah, J. K., Blanz, B., Kopainsky, B., and Schoenberg, W.: An endogenous modelling framework of dietary behavioural change in the fully coupled human-climate FRIDA v2.1 model, Geoscientific Model Development, 18, 5997–6022, https://doi.org/10.5194/gmd-18-5997-2025, 2025.

Rezaei, E. E., Webber, H., Asseng, S., Boote, K., Durand, J. L., Ewert, F., Martre, P., and MacCarthy, D. S.: Climate change impacts on crop yields, Nat Rev Earth Environ, 4, 831–846, https://doi.org/10.1038/s43017-023-00491-0, 2023.

Richardson, G. P.: Reflections on the foundations of system dynamics, System Dynamics Review, 27, 219–243, https://doi.org/10.1002/sdr.462, 2011.

Ristaino, J. B., Anderson, P. K., Bebber, D. P., Brauman, K. A., Cunniffe, N. J., Fedoroff, N. V., Finegold, C., Garrett, K. A., Gilligan, C. A., Jones, C. M., Martin, M. D., MacDonald, G. K., Neenan, P., Records, A., Schmale, D. G., Tateosian, L., and Wei, Q.: The persistent threat of emerging plant disease pandemics to global food security, Proceedings of the National Academy of Sciences, 118, e2022239118, https://doi.org/10.1073/pnas.2022239118, 2021.

Rydzak, F., Obersteiner, M., Kraxner, F., Fritz, S., and McCallum, I.: FeliX3–Impact Assessment Model: Systemic View across Societal Benefit Areas beyond Global Earth Observation, Laxenburg: International Institute for Applied Systems Analysis (IIASA), 2013.

Sang, S., Li, Y., Hou, C., Zi, S., and Lin, H.: The interprovincial green water flow in China and its teleconnected effects on the social economy, Hydrol. Earth Syst. Sci., 29, 67–84, https://doi.org/10.5194/hess-29-67-2025, 2025.

Song, S., Wen, H., Wang, S., Wu, X., Cumming, G. S., and Fu, B.: Quantifying the effects of institutional shifts on water governance in the Yellow River Basin: A social-ecological system perspective, Journal of Hydrology, 629, 130638, https://doi.org/10.1016/j.jhydrol.2024.130638, 2024.

Vaidyanathan, G.: Integrated assessment climate policy models have proven useful, with caveats, Proc. Natl. Acad. Sci. U.S.A., 118, e2101899118, https://doi.org/10.1073/pnas.2101899118, 2021.

Verburg, P. H., Dearing, J. A., Dyke, J. G., Leeuw, S. van der, Seitzinger, S., Steffen, W., and Syvitski, J.: Methods and approaches to modelling the Anthropocene, Global Environmental Change, 39, 328–340, https://doi.org/10.1016/j.gloenvcha.2015.08.007, 2016.

van Vliet, J.: Direct and indirect loss of natural area from urban expansion, Nat Sustain, 2, 755–763, https://doi.org/10.1038/s41893-019-0340-0, 2019.

Waidelich, P., Batibeniz, F., Rising, J., Kikstra, J. S., and Seneviratne, S. I.: Climate damage projections beyond annual temperature, Nat. Clim. Chang., 14, 592–599, https://doi.org/10.1038/s41558-024-01990-8, 2024a.

Waidelich, P., Batibeniz, F., Rising, J., Kikstra, J. S., and Seneviratne, S. I.: Climate damage projections beyond annual temperature, Nat. Clim. Chang., 14, 592–599, https://doi.org/10.1038/s41558-024-01990-8, 2024b.

Wang, T., Teng, F., and Zhang, X.: Assessing global and national economic losses from climate change: a study based on CGEM-IAM in China, Clim. Change Econ., 11, 2041003, https://doi.org/10.1142/S2010007820410031, 2020.

Xu, W., Xiao, Y., Zhang, J., Yang, W., Zhang, L., Hull, V., Wang, Z., Zheng, H., Liu, J., Polasky, S., Jiang, L., Xiao, Y., Shi, X., Rao, E., Lu, F., Wang, X., Daily, G. C., and Ouyang, Z.: Strengthening protected areas for biodiversity and ecosystem services in China, Proc. Natl. Acad. Sci. U.S.A., 114, 1601–1606, https://doi.org/10.1073/pnas.1620503114, 2017.

Yang, Z., Li, Q., Xue, W., and Xu, Z.: Impacts of nature reserves on local residents' income in China, Ecological Economics, 199, 107494, https://doi.org/10.1016/j.ecolecon.2022.107494, 2022.

Ye, Q., Liu, Q., Swamy, D., Gao, L., Moallemi, E. A., Rydzak, F., and Eker, S.: FeliX 2.0: An integrated model of climate, economy, environment, and society interactions, Environmental Modelling & Software, 179, 106121, https://doi.org/10.1016/j.envsoft.2024.106121, 2024.

Zhang, H., Li, R., Cai, X., Zheng, C., Liu, L., Liu, M., Zhang, Q., Lin, H., Chen, L., and Wang, X.: Do electricity flows hamper regional economic–environmental equity?, Applied Energy, 326, 120001, https://doi.org/10.1016/j.apenergy.2022.120001, 2022.

We deeply appreciate the detailed and constructive comments provided by the three anonymous reviewers. Following their suggestions and comments, we have extensively revised the manuscript and provided a point-to-point response to each comment. The original comments are in **bold** font, our response is in regular font, and the changes in the text are in blue.

Comment 2

Sang et al develops a regional system dynamics model to represent coupled human natural systems in the Yellow River Basin, integrating population, economy, energy, food, water, land, carbon, and climate processes to explore historical dynamics and future baseline projections. Overall it looks good, with extensive explaination of the model structure and some illustrative results. However, my main concern is that the human dimension, especially the energy sector, is relatively weakly specified to support the strong claims made about cross sector dynamics in this coupled framework. Because energy use is tightly linked with economic activity, water demand, food production, and carbon emissions, a simplified or poorly validated energy representation risks propagating bias across the entire system. So it's unclear to me whether the projected interactions among energy, water, food, and carbon are internally consistent or robust, especially in the future projections that are central to the paper's conclusions.

Response: We appreciate the reviewer's concern regarding the energy sector specification. We acknowledge that the energy module is simplified to balance overall system complexity; however, we have verified its robustness by extending the validation with recent data (2021–2024), which demonstrates a good match with extended historical data. Furthermore, since the energy sector accounts for a relatively minor fraction of total basin-wide water consumption compared to agriculture (>70%), this simplification does not propagate significant bias to the study's core human–water interaction conclusions. We have refined the description of the energy logic and explicitly discussed limitations, such as electrification trends, as priorities for future model iterations in the revised manuscript.

1. **First, A key issue is the system boundary of the energy sector. Although an energy module is included, the model appears to explicitly represent only electricity, while other major energy uses, particularly in industry, are treated in a highly aggregated manner. This is problematic because non electric energy use and structural change in industry are major drivers of both emissions and water demand. In addition, important indirect effects are not clearly represented. For example, electrification trends such as electric vehicle deployment would increase electricity demand while reducing oil consumption, yet such substitution dynamics are not discussed. Similarly, changes in industrial energy structure could feed back to economic output and water use, but it is unclear how these interactions are captured. Given the stated goal of representing coupled human natural systems, it is important for the authors to clarify how these cross sector linkages are treated and how sensitive the results are to the simplified energy boundary.**

Response: We acknowledge that compared to dedicated bottom-up energy models, such as those included in IAMs, the energy sector in the CHANS-SD-YRB 1.0 model has been simplified to balance the complexity of the coupled human-natural system in the YRB. Given the breadth of human-nature interactions in the study region, a model could not exhaust every interaction and pathway. Therefore, the current model focuses on simulating the basin's human–water interactions through water supply and demand perspectives, with necessary simplifications made for certain processes and sectors. We completely agree that electrification (e.g., EV deployment) and AI data center are critical emerging trends for energy consumption growth in recent years. While electrification alters the energy mix, its direct impact on the water balance of the Yellow River Basin is of secondary compared to agricultural and ecological water demands. We fully recognize the importance and incorporating these mechanisms into a more detailed energy substitution module (reflecting trends like EV adoption) can expand model's capability for simulating finer-scale energy-water nexus dynamics, which we will consider for the next major release of the model.

The reviewer correctly points out that industry and energy structure affect water demand. In the Yellow River Basin, water withdrawal is overwhelmingly dominated by agriculture (>70%) whereas industrial water withdrawal only accounts for 12% in 2024. Although thermal power is the most water-intensive in energy source sector, its share in total basin-wide water consumption is relatively small, and in many years, even zero. (data from China Water Resources Bulletin http://www.mwr.gov.cn/sj/#tjgb). Moreover, there is a lack of long-term data on specific industry sectors and their water withdrawal, making explicitly modeling industry sectors and their structural changes difficult. Given the data constraint, the aggregated approach for the energy structure is a practical design choice to satisfy our modeling goal to capture the main driver of water pressure.

In the current model, the feedback loop from changing energy structure to economic output and water use is not explicitly represented. The model employs a demand-driven logic for the energy sector, where energy supply is assumed to meet the demand generated by economic activities. Therefore, we do not currently simulate scenarios where energy shortages or structural constraints restrict GDP growth.

In response to these comments, we have incorporated a new subsection into the Discussion. The revised texts are shown below:

For human processes, the Energy and Economy sectors could be refined to model more detailed industry subsectors and emerging trends in energy demand driven by electrification.

2. **Second, the description of the energy sector in Section 2.2.3 is unclear. The statement that "coal, oil, and gas consumption are derived from linear relationships between historical sectoral GDP" and corresponding consumption is ambiguous. It is not clear whether this refers to total final energy consumption or only to fuels used for electricity generation. More**

importantly, a linear extrapolation based on historical GDP trends is unlikely to be appropriate for long term projections under ongoing structural change, efficiency improvements, and decarbonization. The manuscript also does not clearly explain how total energy demand evolves over time or how energy intensity changes are represented. Besides, electricity generation shares are imposed exogenously based on Li et al. 2024, which makes the future pathway highly deterministic. The underlying policy assumptions and narratives are not clearly described. So Figure 16 would benefit from validation against recent historical data before 2025 and comparison with ranges reported in the literature. In particular, the emerging nuclear capacity around 2030 requires stronger justification in terms of feasibility and policy assumptions in China's context.

Response: We apologize for the lack of clarity in the description of the *Energy* sector. As for the ambiguity regarding coal, oil, and gas consumption, these fossil fuels refer to consumption by production of industry and service sectors, which is derived from their relationship with sectoral GDP. Fossil fuel consumption for electricity generation by thermal power is separately calculated based on total electricity consumption and the generation mix. These two constitutes the total coal, oil, and gas consumption.

We have significantly revised the description of the energy sector in Section 2.2.3 and provided detailed description in the Supporting Information Section S2.4:

The fossil fuel consumption by economic production of industry and service sectors is modeled as a linear function of sectoral GDP based on historical data (Equations S25-S27).

The detailed information in the Supporting Information is shown below:

Coal (Equation S25), oil (Equation S26) and gas (Equation S27) consumption by economic production of industry and service sectors are obtained from the linear fit of historical sector GDP and related consumption,

$$Coal_{con} = TP \times Share_{coal} \times Coef_{ESC} \times Coef_{SCC} + Para_{IC} \times GDP_{ind} + Const_{IC} \quad \text{(S25)}$$

$$Oil_{con} = TP \times Share_{oil} \times Coef_{ESC} \times Coef_{SCO} + Para_{ISO} \times (GDP_{ind} + GDP_{ser}) + Const_{ISO} \quad (S26)$$

$$Gas_{con} = TP \times Share_{gas} \times Coef_{ESC} \times Coef_{SCG} + Para_{ISG} \times (GDP_{ind} + GDP_{ser}) + Const_{ISG} + Pop_{rural} \times PCG_{rural} + Pop_{urban} \times PCG_{urban} \quad (S27)$$

where *TP* is the thermal power generation; $Share_{coal}$, $Share_{oil}$ and $Share_{gas}$ are the proportion of coal, oil and gas in thermal power generation; $Coef_{ESC}$ is the standard coal conversion coefficient of electricity, $Coef_{SCC}$, $Coef_{SCO}$ and $Coef_{SCG}$ are the conversion coefficients of standard coal, coal, oil and gas, which are derived from General Principles for the Calculation of Comprehensive Energy Consumption (GB/T 2589-2008, GB/T 2589-2020); $Para_{IC}$ and $Const_{IC}$ represent the slope and intercept derived from the linear fit of historical industry GDP and coal consumption. Similarly, $Para_{ISO}$, $Const_{ISO}$ and $Para_{ISG}$, $Const_{ISG}$ denote the corresponding parameters for oil and gas, respectively. Notably, these parameters were estimated using a piecewise linear regression with distinct coefficients for the pre- and post-2010 periods to capture the structural changes around 2010 driven by socio-economic shifts. Pop_{rural} and Pop_{urban} are rural and urban population, from *Population* sector, PCG_{rural} and PCG_{urban} are the per capita natural gas consumption in rural and urban areas from the China Energy Statistical Yearbook (NBSC, 2020a).

Regarding the concern about "linear extrapolation," we acknowledge that a single linear trend cannot capture long-term decarbonization. Future decarbonization scenario can be designed by adjusting this parameter. For historical period, we employed a piecewise linear regression approach with a structural break around 2010. The change in slope after 2010 explicitly captures the decoupling trend between economic growth and energy consumption (i.e., efficiency improvements and structural change).

Total energy demand is driven by socio-economic development. Electricity demand is calculated based on sectoral electricity intensities, while fossil fuel demand (coal, oil, gas) is derived from linear relationships with sectoral GDP, where the coefficients serve as proxies for the marginal energy intensity of production activities. These parameters are calibrated at the provincial level to capture spatially divergent trends.

Since future energy structure changes are uncertain, we used electricity generation shares projection following the carbon neutrality goals of China by Li et al. (2024) in the future baseline scenario. China’s energy transition is heavily driven by top-down state planning (i.e., Carbon Peaking and Carbon Neutrality Goals), modeling these shares as exogenous policy targets allows us to simulate the specific impact of fulfilling national strategic goals, rather than relying on a market-clearing mechanism which might misrepresent China's regulated energy sector.

Regarding the validation period, we found that electricity generation data is available up to 2024, while fossil fuel consumption data (coal, oil, and natural gas) is currently only published up to 2022. We have collected these latest datasets and compare our simulation results against this extended historical period (2021–2024 for electricity, 2021–2022 for fossil fuels). Please refer to the response to the next question for the figure.

3. **Third, the paper emphasizes the regional and spatially explicit nature of the CHANS SD YRB model, but this advantage is not fully demonstrated in the results. Although Figure 1 highlights spatial resolution across provinces and sub basins, many subsequent figures are conceptual diagrams rather than quantitative outputs. Model validation for energy variables ends in 2020, and maybe there are more recent to validate, such as 2021-2024? It would be important to show how near term projections after 2020 compare with observations, particularly for electricity generation and fuel consumption. Also, author may consider presenting and validating provincial level energy results beyond more aggregated models.**

Response： Thank you for the comments.

We would like to clarify that the regional and spatially explicit nature is a core feature of our model and the simulation results. The results are not merely aggregated; they maintain spatial resolution throughout the simulation. As shown in the "Results" section, the human sectors’ outputs (e.g., economic and population metrics) are presented for the nine provinces, while the natural sectors' outputs (e.g., runoff) are validated and projected at the sub-basin level (up-, mid-, and downstream). The future scenario results explicitly display trajectories for both the aggregate Yellow River Basin and the individual nine provinces.

Regarding the validation period, this study was initiated in 2022. At that time, the official statistical yearbooks for all nine provinces were only available up to 2020 due to the standard 2-year lag in regional data publication in China. While some fragmented national data for 2021–2024 exist, integrating inconsistent datasets would compromise the model's internal coherence. Therefore, we maintained 2020 as the uniform cutoff for calibration to ensure data consistency across all sectors.

In response to the reviewer's suggestion, we attempted to collect more recent data for *Energy* sector covering the period 2021–2024. However, the official statistics are currently limited. The data for electricity generation are available for the full 2021–2024 period, consumption data for coal, oil, and natural gas are only available for 2021–2022. These available historical statistical data after 2020 are added in the comparison with the simulation results in Figure 1. The simulation results show good agreement with the extended historical data in terms of both trends and magnitudes.

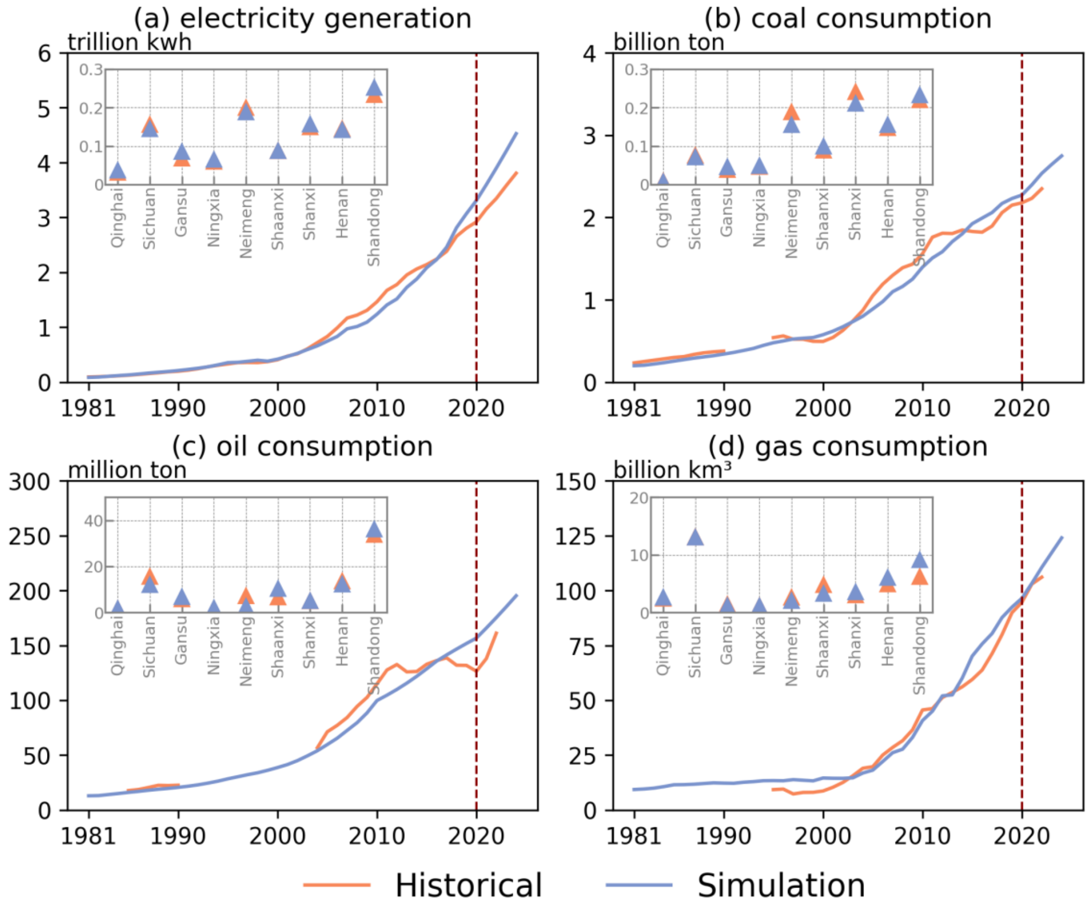

Fig.1 Validation of variables in *Energy* sector during historical period of 1981-2024. Pink triangles in the upper left sub-image represents the average of historical and simulation value in 1981-2024 of each province in the YRB.

We agree that provincial validation is crucial, which is included in the inset of Figure 1. The triangular markers represent the comparison between the simulated means and the historical statistical data for each of the nine provinces. This visualization explicitly shows the model's performance at the provincial level, demonstrating its capability to capture regional variations beyond the basin level.

4. **Lastly, there are several clarity issues that should be addressed. The meaning of "fire" in the energy sector in Figure 2 is unclear. In addition, the term "per capital" appears repeatedly in the text and figures and should be corrected to "per capita" throughout the manuscript.**

Response: Thank you for the comments.

We apologize for the typographical errors in the manuscript. We have corrected them in the revised version. Specifically, the label 'fire' in Figure 2 has been updated to 'Thermal Power', and all instances of 'per capital' have been corrected to 'per capita' throughout the text.

The revised Figure 2 is shown below:

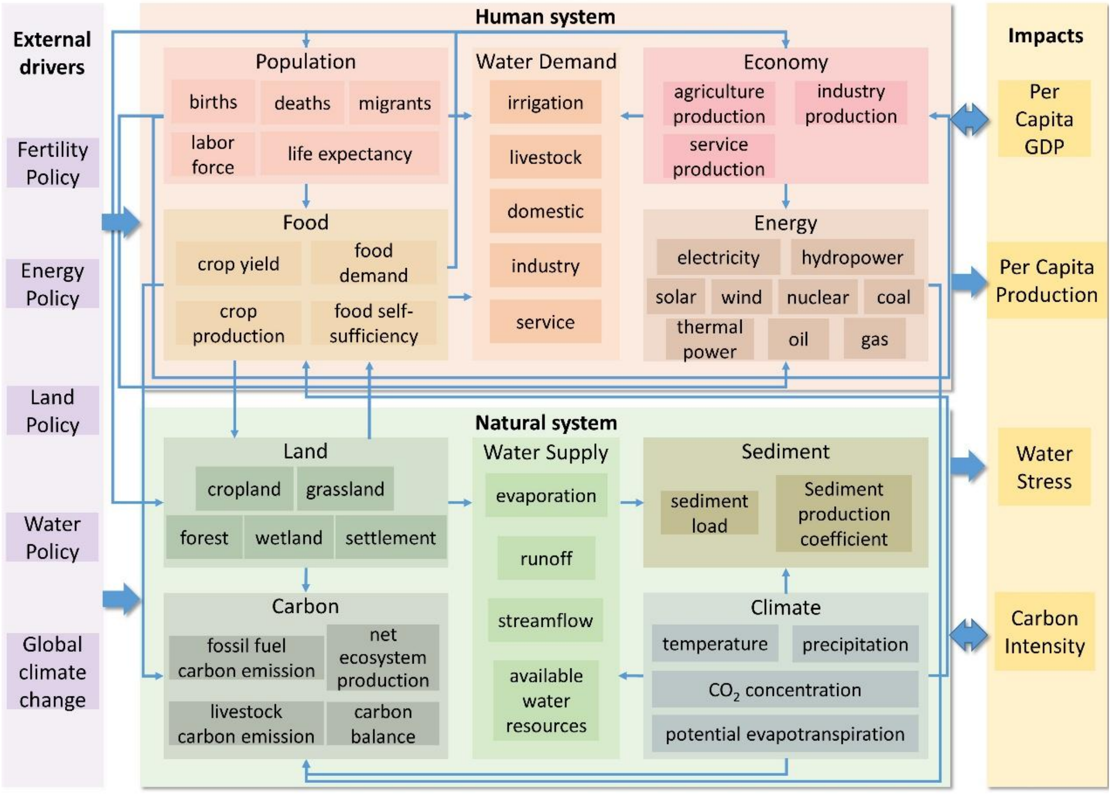

Fig.2 (Fig.2) Structure of the CHANS-SD-YRB, which shows sectors of human and natural systems, their key processes and interactions.

References:

Li, M., Shan, R., Abdulla, A., Virguez, E., and Gao, S.: The role of dispatchability in China's power system decarbonization, Energy Environ. Sci., 17, 2193–2205, https://doi.org/10.1039/D3EE04293F, 2024.

National Bureau of Statistics of China (NBSC), China Energy Statistical Yearbook, https://data.cnki.net/yearBook?type=type&code=A, 2020.

We deeply appreciate the detailed and constructive comments provided by the three anonymous reviewers. Following their suggestions and comments, we have extensively revised the manuscript and provided a point-to-point response to each comment. The original comments are in **bold** font, our response is in regular font, and the changes in the text are in blue.

Comment 3

This manuscript develops a regional coupled human–natural systems model for the Yellow River basin (CHANS SD-YRB) using a System Dynamics framework that integrates 10 interacting sectors at annual, provincial, and sub-basin scales while explicitly conserving hydrological connectivity among sub-basins. The model reproduces historical dynamics with strong reported agreement and is then used for baseline projections under continued policies and SSP245. The key novelty is the basin-specific, multi-sector CHANS platform that couples socio-economic dynamics with water, sediment, land, carbon, and climate processes in a single connected sub-basin framework, enabling diagnosis of cross-sector feedbacks and trade-offs relevant for long-term sustainability planning. However, some revisions are required before the manuscript can be considered for publication.

Response: Thank you for taking your time to review our study and provide feedback and comments. Following the suggestions and comments, we have extensively revised the manuscript and provided a point-to-point response to each comment.

1. **Could the authors clarify the key advantages of this study (and the CHANS SD-YRB model) compared with Sang et al. (2025)?**

Response： We thank the reviewer for giving us the opportunity to clarify the relationship and distinct contributions of this study compared to Sang et al. (2025).

The conceptual modeling framework for the Yellow River Basin CHANS proposed by

Sang et al. (2025) is a theoretical architecture with identified key feedback loops to guide system model integration. This study implements the theoretical modeling framework into a fully operational and open-source System Dynamics model CHANS-SD-YRB 1.0, including the detailed mathematical equations, parameter values, and internal structures for all coupled sectors. With rigorous historical calibration and validation, we demonstrate the feasibility and accuracy of the model to project future CHANS dynamics, providing specific quantitative evidence to support the practical suggestions outlined in the conceptual framework.

We have added the clarification to the last paragraph in the manuscript. The revised texts are shown below:

Drawing on the conceptual framework of Sang et al. (2025), this study implements it into a fully functional, validated System Dynamics model tool, CHANS-SD-YRB 1.0.

2. **Figure 4: The model appears not to include a price response to market changes. Under scenarios with carbon pricing, this assumption may miss some impacts of the carbon price. Could the authors discuss this limitation?**

Response: Thank you for pointing out this issue.

The current CHANS SD-YRB model does not include an endogenous price-clearing mechanism but uses exogenous prices. Following the reviewer's suggestion, we have added a dedicated paragraph in the Discussion section to articulate this limitation. We clarify that while the model captures the physical results of climate policies, it cannot evaluate the economic efficiency of specific pricing mechanisms. The revised texts are shown below:

Moreover, there are still important feedbacks absent from the current coupling framework. Notable examples include the effects of land use change on climate, the effects of climate change on economic growth, and the effects of pricing on energy use and carbon emissions. These missing feedbacks could be incorporated based on recent studies, including the land use feedback on precipitation through moisture recycling (Sang et al., 2025), socioeconomic losses from climate change impacts (Waidelich et al., 2024), and integration with Computable General Equilibrium (CGE) models (Fujimori et al., 2014).

3. **Lines 284: Electricity generation shares are treated as exogenous and sourced from the yearbook. Do these shares account for future changes? If not, given that the SSP2 baseline is above 6.0 W/m² while this study uses SSP245 (which still implies future climate policy), would assuming fixed shares remain reasonable?**

Response: Thank you for the comments.

We apologize for the ambiguity. The electricity shares are not fixed but time-varying for both historical and future projection. While historical shares are sourced from statistical yearbooks, the future shares in the Baseline Scenario are sourced from the projections by Li et al. (2024) including year-by-year projections (2021-2060) of the generation mix (Coal, Hydro, Wind, Solar, Nuclear) for each province to meet China's national Carbon Peaking and Carbon Neutrality goals.

4. **Line 92: Could the authors please double-check the citation format here?**

Response: Thank you for the comments.

The variation in the citation format (including the initial "X.") was automatically generated by our citation management software (Zotero) to disambiguate two different first authors with the surname "Li" published in the same year.

To ensure consistency with the journal's citation style, we have manually corrected this in the revised manuscript. We now use the standard suffix format (e.g., Li et al., 2018a; Li et al., 2018b) to distinguish these references without including the first initial.

5. **Line 178: Could the authors clarify how these variables are used as proxies for disaggregation?**

Response: Thank you for the comments.

The gridded data serves as the basis for calculating spatial weighting factors to downscale the model's provincial-level outputs to the sub-basin scale (up-, mid-, and downstream). We utilized high-resolution population grid data to calculate the proportion of each province's population located within the boundaries of the up-, mid-, and downstream. These calculated weights were then used to convert provincial-level demographic variables (e.g., total population, labor force) into sub-basin scale. Similarly, we employed GDP grid data to determine the spatial distribution of economic activity. The derived weights (the share of a province's GDP falling into each sub-basin) were used to disaggregate economic and emission-related variables, specifically GDP and human carbon emissions from the provincial scale to the basin scale.

In the revised manuscript and Supporting Information, we have clarified. The revised texts are shown below:

For manuscript: Given the availability of gridded data for human processes and the strong correlations among relevant variables, gridded population and GDP data were used as proxies to disaggregate demographic variables and economic and human carbon emissions, respectively, from provincial-level to the sub-basin scale (Table S3).

For Supporting Information: Using high-resolution gridded data, we calculated the spatial weights of each province within the up-, mid-, and downstream to convert provincial to sub-basin values.

The spatial weights derived from gridded population were used as proxies to disaggregate demographic outputs (e.g., total population and food demand), while weights derived from gridded GDP were applied to disaggregate economic and environmental outputs (e.g., GDP and anthropogenic carbon emissions) from the provincial level to the sub-basin scale.

6. **Equations (5) and (6): Could the authors explain why the functions are specified in this form?**

Response： Thank you for the comments.

The functional forms were chosen based on the distinct physical mechanisms and socio-economic development rules governing different water use sectors.

For Equation (5), it specifies a process-based physical coupling mechanism between the *Land* and *Water Demand* sectors.

For Equation (6), it employs a non-linear saturation function rather than a simple linear regression. Empirical studies (Flörke et al., 2013) suggest that per capita water use rises rapidly during early stages of economic development (due to improved sanitation and appliance usage) but eventually stabilizes or saturates at high income levels.

We have added these explanations in the revised manuscript. The revised texts are shown below:

Irrigation water withdrawal (WW_{irr}) is estimated through a physically-based function of exogenous irrigation water use intensity (WWI_{irr}), cropland irrigation ratios (IR) from the China Agricultural Yearbook (MAARA, 2020), and cropland area ($Area_{Cropland}$) provided by the *Land* sector (Equation 5).

Residential water withdrawal (WW_{res}) is derived from a non-linear empirical function of economic development and population growth with an upper limit of per capita domestic water use (Flörke et al., 2013) because water demand per person cannot increase indefinitely (Equation 6).

7. Lines 467: Could the authors elaborate on why SSP245 aligns more closely with the climate trends in the YRB?

Response: Thank you for the comments.

For the future baseline scenario presented, we selected SSP2-4.5 because it represents a "middle-of-the-road" trajectory that align with the ongoing real-world climate change. This trajectory aligns closely with China's current national strategy, specifically the goal to peak carbon emissions before 2030 and achieve deep reductions thereafter. In contrast, SSP5-8.5 assumes fossil-fueled development which contradicts current policies, while SSP1-2.6 assumes an immediate and drastic sustainability transition that may be overly optimistic for the developing western regions of the YRB. So we consider SSP2-4.5 to be the most representative baseline for the Yellow River Basin.

References:

Flörke, M., Kynast, E., Bärlund, I., Eisner, S., Wimmer, F., and Alcamo, J.: Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development: A global simulation study, Global Environmental Change, 23, 144–156, https://doi.org/10.1016/j.gloenvcha.2012.10.018, 2013.

Fujimori, S., Masui, T., and Matsuoka, Y.: Development of a global computable general equilibrium model coupled with detailed energy end-use technology, Applied Energy, 128, 296–306, https://doi.org/10.1016/j.apenergy.2014.04.074, 2014.

Ministry of Agriculture and Rural Affairs of the People's Republic of China (MAARA), China Agriculture Yearbook, https://doi.org/https://data.cnki.net/yearBook?type=type&code=A, 2020.

Sang, S., Li, Y., Hou, C., Zi, S., and Lin, H.: The interprovincial green water flow in China and its teleconnected effects on the social economy, Hydrol. Earth Syst. Sci., 29, 67–84, https://doi.org/10.5194/hess-29-67-2025, 2025.

Waidelich, P., Batibeniz, F., Rising, J., Kikstra, J. S., and Seneviratne, S. I.: Climate damage projections beyond annual temperature, Nat. Clim. Chang., 14, 592–599, https://doi.org/10.1038/s41558-024-01990-8, 2024.