Metadata Driven Modeling of Green Finance Transmission Paths Influencing Economic Growth in Ethnic Mountainous Regions

Ke Chen[1]
[1]School of Management Engineering, Chengdu Technological University, Chengdu, Sichuan, 611730, China
Corresponding author: Ke Chen (e-mail: 13086606571@163.com).

ABSTRACT To address the challenges of aligning heterogeneous data from multiple sources in ethnic minority mountainous areas and the difficulty in identifying the nonlinear time-delay structure of green finance transmission paths, this paper first constructs a heterogeneous graph containing five types of nodes and three types of edges using metadata. Then, a dynamic graph variational encoder is used to output the latent variables of candidate paths. Finally, a gated cyclic decoder is used to simulate the time-delay transmission and attenuation of funds, jointly optimizing prediction error, path sparsity regularization, and causal prior constraints. Experiments identified six core transmission paths with a prediction determination coefficient of 0.865, and all path coefficients in the posterior interval did not contain zeros. The results demonstrate that the metadata-driven graph variational autoencoder can discover interpretable transmission paths end-to-end, providing a quantitative basis for green finance policies in ethnic minority mountainous areas.

INDEX TERMS Green Finance; Transmission Path; Economic Growth; Ethnic Minority Mountainous Areas; Dynamic Graph Variational Autoencoder

I. INTRODUCTION

DUE to the complex terrain, ecological fragility, and underdeveloped financial markets in ethnic minority mountainous areas, the transmission path of green finance, as a key tool for promoting the synergy between ecological protection and economic growth, is still unclear [1–2]. Existing studies mostly use panel regression or structural equation models, but when faced with multi-source heterogeneous data (such as financial project records, remote sensing geographic grids, and ethnic population statistics), it is difficult to automatically align information from different modalities and capture the transmission structure containing nonlinear time delays, resulting in a lack of interpretability in path decomposition and limiting the accurate formulation of green finance policies [3].

To address the aforementioned issues, existing research has begun to utilize graph neural networks to characterize the relational structures in financial systems, multivariate time series, and spatial networks. Wang et al. pointed out in their review of financial graph neural network research that financial data often exists in the form of graph structures such as subjects, assets, transactions, and relationships [4]. Bloemheuvel et al. applied graph neural networks to multivariate time series regression, demonstrating that this type of method can characterize time dynamics while introducing structural relationships between variables [5]. Jin et al. systematically summarized the applications of graph neural networks in time series prediction, classification, imputation, and anomaly detection [6]. Wang's review of spatiotemporal graph neural network prediction research shows that this type of model can provide an effective framework for spatial correlation and time dynamics modeling in spatiotemporal data prediction [7]. Ye et al. explained from the perspective of knowledge graphs that graph neural networks have been widely used in tasks such as knowledge graph link prediction, graph alignment, relation reasoning, and node classification [8]. The above studies provide a methodological foundation for modeling the transmission path of green finance. However, for the special scenario of ethnic minority mountainous areas, which has the characteristics of terrain barriers, differences in ethnic settlements and uneven financial accessibility, existing models still rarely incorporate prior information such as geographical altitude, ethnic distribution density, spatial distance and policy time lag into the graph structure learning process. As a result, the generalization ability and path explanation ability of the models across different sub-regions are still limited.

In contrast, metadata-driven methods can organize multi-source heterogeneous data into a computable graph structure through semantic alignment, rule extraction, and structured mapping. Masmoudi et al. pointed out in their research on semantic data integration that semantic description, metadata constraints, and query mechanisms can improve the consistent expression and cross-source association capabilities among heterogeneous data [9]. Zhong et al.'s review on automatic knowledge graph construction further shows that processes such as entity recognition, relation extraction, knowledge acquisition, knowledge fusion, and knowledge evolution can transform scattered data into a structured knowledge network with explicit semantic relationships [10]. Hofer et al. emphasized that high-quality knowledge graph construction not only involves structured and unstructured data processing, but also requires supporting mechanisms such as metadata management, ontology development, and quality assurance [11]. Inspired by the aforementioned research, this paper employs a metadata-driven dynamic graph variational autoencoder to map green finance projects, industry sectors, geographic rasters, and ethnic settlements as nodes in a heterogeneous graph. Edge features are defined based on semantic similarity, spatial distance, and temporal lag coefficients from the metadata. The encoder outputs latent variables for candidate transmission paths through metadata attention graph convolution, while the decoder uses gated recurrent units to simulate the time-lag propagation and attenuation of capital investment along each path. The algorithm jointly optimizes prediction error, path sparsity regularization, and causal direction prior constraints. This paper aims to automatically identify the core transmission paths by which green finance impacts economic growth in ethnic minority mountainous areas from multi-source data and quantify the strength and time-lag characteristics of different paths.

II. METHOD

A. CONSTRUCTION OF METADATA HETEROGENEOUS GRAPH

The construction of the heterogeneous metadata map begins with raw, multi-source data. This raw data includes records of green finance projects in ethnic minority mountainous areas, statistics from economic and industrial sectors, remote sensing geographic rasters, distribution of ethnic settlements, and corresponding metadata description files. The metadata description files define the semantic labels, spatial reference frame, temporal granularity, missing value markers, and predefined causal priors for each data field. Based on this metadata, all entities are mapped to nodes in the map. Node types are categorized into five types: green finance nodes (e.g., specific loan products, green bonds), industrial sector nodes (e.g., ecotourism, specialty planting, clean energy), geographic raster nodes (spatial units divided by altitude and slope), ethnic settlement nodes (residential areas identified by ethnic composition), and time nodes (years or quarters). Each node is accompanied by a static attribute vector provided by the metadata; for example, the attributes of a green finance node include total funding amount, term, and interest rate type encoding; the attributes of a geographic raster node include average altitude, slope, and annual precipitation.

The connection between nodes is driven by the association rules recorded in the metadata. For any pair of nodes, if there is a direct association description in the metadata, an edge is established between the pair of nodes. Each edge is assigned an initial feature vector, which is composed of three components: semantic similarity component, spatial distance component, and time lag coefficient component [12]. Let the heterogeneous graph be $G = (V, \mathcal{E}, X, E_{\text{feat}})$, where V is the set of nodes, $\mathcal{E} \subseteq V \times V$ is the set of edges, $X \in \mathbb{R}^{|V| \times d_v}$ is the node attribute matrix, $E_{\text{feat}} \in \mathbb{R}^{|\mathcal{E}| \times d_e}$ is the edge feature matrix, d_v is the node attribute dimension, and d_e is the edge feature dimension. In formula (1), for any edge $(u, v) \in \mathcal{E}$, its edge feature vector e_{uv} is defined as (1):

$$e_{uv} = [\text{sim}_{\text{sem}}(u, v);\ \alpha_{uv} \cdot d_{\text{geo}}(u, v);\ \tau_{uv}]. \tag{1}$$

In formula (1), $\text{sim}_{\text{sem}}(u, v)$ is the semantic embedding similarity of the metadata between nodes u and v, with a value between 0 and 1; $d_{\text{geo}}(u, v)$ is the spatial distance between nodes (set to 0 if the node does not involve spatial location); α_{uv} is the spatial weight scaling factor provided in the metadata, reflecting the degree to which terrain hinders economic connections; τ_{uv} is the preset initial time lag factor of the metadata, in years. All edge feature vectors constitute the matrix E_{feat}. Figure 1 shows the complete data flow of the above construction process.

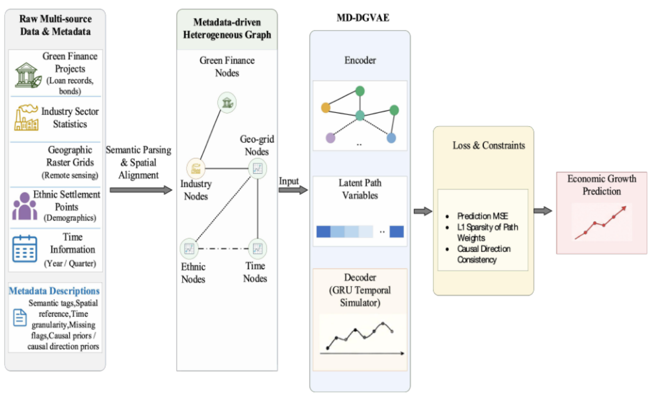

FIGURE 1. Metadata heterogeneity graph and MD-DGVAE data flow framework diagram.

In Figure 1, the left side shows the original multi-source data and its metadata description. After semantic parsing and spatial alignment, the data converges into a heterogeneous graph in the middle section. The graph contains five types of nodes and three types of edges. The right side of the graph connects to the graph convolutional layer of the encoder. Meanwhile, the attention weights and causal priors in the metadata are indicated by dashed arrows pointing to the aggregation module of the graph convolution and the loss function of the subsequent decoder, respectively.

B. DYNAMIC GRAPH VARIATIONAL ENCODER

The dynamic graph variational encoder receives the heterogeneous graph $G = (V, \mathcal{E}, X, E_{\text{feat}})$ constructed in the previous section as input and outputs the latent variable distribution of candidate propagation paths. The encoder consists of multiple stacked graph convolutions. Each layer introduces a metadata attention mechanism during graph convolution to dynamically adjust the aggregation weights between different neighboring nodes. This attention weight is derived from the fusion of the semantic similarity component $\text{sim}_{\text{sem}}(u, v)$ and the spatial distance component $\alpha_{uv} \cdot d_{\text{geo}}(u, v)$ in the edge feature vector e_{uv}. The original attention coefficient after fusion is $\omega_{uv}^{(0)} = \text{sim}_{\text{sem}}(u, v) \cdot \exp(-\beta \cdot \alpha_{uv} d_{\text{geo}}(u, v))$, where β is the spatial attenuation coefficient, determined by the terrain impedance factor extracted from the metadata. For node u, the aggregation weight $\omega_{uv}^{(l)}$ of the l-th layer is calculated after normalization of the attention coefficients of all its outgoing neighbors.

$$\omega_{uv}^{(l)} = \frac{\exp\left(\text{LeakyReLU}\left(\mathbf{a}^{(l)\top}\begin{bmatrix} W^{(l)}h_u^{(l-1)} \| W^{(l)}h_v^{(l-1)} \\ \| E_{\text{feat}}^{(uv)} \end{bmatrix}\right)\right)}{\sum_{k \in N(u)} \exp\left(\text{LeakyReLU}\left(\mathbf{a}^{(l)\top}\begin{bmatrix} W^{(l)}h_u^{(l-1)} \| W^{(l)}h_k^{(l-1)} \\ \| E_{\text{feat}}^{(uk)} \end{bmatrix}\right)\right)}. \tag{2}$$

In formula (2), $h_u^{(l-1)}$ is the hidden state of node u in layer $l-1$, $W^{(l)}$ is the learnable weight matrix of layer l, $E_{\text{feat}}^{(uv)}$ is the complete feature vector of edge (u, v), $\mathbf{a}^{(l)}$ is the attention parameter vector of layer l, $\|$ represents vector concatenation, $N(u)$ is the set of neighboring nodes of node u, and LeakyReLU is the non-linear activation function. The normalized attention weights are used for weighted aggregation of the neighboring hidden states to obtain the updated state of node u in layer l:

$$h_u^{(l)} = \sigma\left(\sum_{v \in N(u)} \omega_{uv}^{(l)} W^{(l)} h_v^{(l-1)}\right). \tag{3}$$

In formula (3), σ is the ELU activation function. After L layers of graph convolution, the final hidden state $h_u^{(L)}$ of each node incorporates the local graph structure information guided by metadata [13]. The encoder aggregates the hidden states of all nodes into a global graph representation vector z_{graph}, and uses the readout function to perform a weighted summation of the hidden states of all nodes in the heterogeneous graph. The weights are determined by the prior importance coefficient γ_t of the metadata corresponding to the node type (the highest value is taken for green finance nodes, the second highest value is taken for ethnic nodes, and the lowest value is taken for geographic raster nodes):

$$z_{\text{graph}} = \sum_{t \in \text{types}} \gamma_t \cdot \sum_{u \in V_t} h_u^{(L)}. \tag{4}$$

In formula (4), V_t is the set of nodes of type t. The global graph representation vector z_{graph} is input into two independent fully connected layers, which output the mean vector μ and the log-variance vector $\log \sigma^2$ of the latent path variable distribution, respectively. The latent variable $z \in \mathbb{R}^{d_z}$ is obtained through reparameterized sampling:

$$z = \mu + \sigma \odot \epsilon, \quad \epsilon \sim N(0, I). \tag{5}$$

In formula (5), $\odot$ represents element-wise multiplication, and ϵ represents standard Gaussian noise. Each dimension of z corresponds to the implicit representation of a candidate propagation path, and its dimension d_z is determined by the maximum number of possible paths preset in the metadata, which is compressed during model training through subsequent sparse regularization. The latent variable z finally output by the encoder is passed to the decoder for reconstructing economic growth prediction.

C. GATED CYCLIC SEQUENTIAL DECODER AND PATH SPARSITY REGULARIZATION

The gated recurrent time series decoder receives the latent variable $z \in \mathbb{R}^{d_z}$ from the encoder, treating each dimension as a latent representation of a candidate transmission path. It then simulates the transmission, decay, and accumulation of green finance inputs along each path over time using a Gated Recurrent Unit (GRU) network, ultimately reconstructing the time series of economic growth indicators for ethnic minority mountainous areas. The decoder's input includes an initial green finance input vector $f_0 \in \mathbb{R}^{d_f}$, where the dimension d_f corresponds to the number of green finance nodes, and each element represents the funding amount of the corresponding financial project at the initial time. The latent variable z is first transformed using a linear transformation to generate a path transition matrix $T \in \mathbb{R}^{d_f \times d_z}$ and a path decay vector $\rho \in \mathbb{R}^{d_z}$, where the element in the i-th row and j-th column of T represents the proportion of input received from the i-th green finance project along the j-th path, and ρ_j is the decay rate at each time step along the j-th path. At each time step t ($t = 1, 2, \ldots, T_{\max}$), the current propagation quantity $p_j^{(t)}$ on path j is updated as follows:

$$p_j^{(t)} = \rho_j \cdot p_j^{(t-1)} + \sum_{i=1}^{d_f} T_{ij} \cdot f_i^{(t-1)}. \tag{6}$$

In formula (6), $p_j^{(t-1)}$ is the stock on path j at the previous time step, and $f_i^{(t-1)}$ is the actual investment amount of the i-th green finance project at the previous time step. For the initial time step $t = 1$, $p_j^{(0)} = 0$, the value of $f_i^{(0)}$ is the original input data. The investment amount $f_i^{(t)}$ of each green finance project at time step t is determined by external policy drivers and path feedback. However, in order to focus on the transmission path itself, it is assumed that $f_i^{(t)}$ remains constant within the time window (that is, $f_i^{(t)} = f_i^{(0)}$). The transfer quantities on all paths are gathered into vectors $p^{(t)} \in \mathbb{R}^{d_z}$, and the marginal contribution to economic growth at this time step $g^{(t)} = Mp^{(t)}$ is obtained through a learnable output mapping matrix $M \in \mathbb{R}^{1\times d_z}$. The decoder uses a single-layer GRU to integrate the marginal contribution series at each time step into the final economic growth forecast. The GRU hidden state $s^{(t)}$ update rule is:

$$s^{(t)} = \mathrm{GRU}\left(s^{(t-1)}, g^{(t)}\right). \tag{7}$$

In formula (7), the initial hidden state $s^{(0)} = 0$. The hidden state $s^{(T_{\max})}$ of the last time step $T_{\max}$ is input into a fully connected layer, and the predicted value $\hat{y} \in \mathbb{R}$ of the economic growth indicator is output.

Path sparsity regularization is applied to the set of path weights related to the latent variable z during the training of the decoder. Specifically, the comprehensive importance coefficient $\lambda_j = \|T_{:,j}\|_2 \cdot \rho_j^{-1}$ is defined for each candidate path j, where $T_{:,j}$ is the j-th column of the transition matrix, and ρ_j^{-1} reflects the persistence of the path. The regularization term $\mathcal{R} = \sum_{j=1}^{d_z} |\lambda_j|$, i.e., the L1 norm, participates in the optimization together with the prediction loss. When λ_j is compressed to zero, the corresponding path is removed from the model, and its latent representation dimension is not included in the calculation in subsequent iterations. Table 1 lists the metadata statistical characteristics of the input data used to construct the decoder for each sub-region of the ethnic minority mountainous area, including altitude range, ethnic distribution density, number of green finance projects, and the original data missing rate of each variable. These statistics directly affect the initialization boundary of the path decay rate ρ and the selection of the time step $T_{\max}$.

TABLE 1. Statistical Characteristics of Metadata in Sub-Regions of Ethnic Minority Mountainous Areas

Sub-region name	Altitude range (m)	Ethnic population density (persons/km^2)	Number of green finance projects	Data missing rate (%)
Valley plain area	800–1500	42.3	31	8.2
Low mountain and hilly area	1500–2200	28.7	24	15.6
High half-mountain area	2200–3000	16.5	15	27.4
Alpine gorge area	3000–4200	6.8	8	41.3

Table 1 shows that the higher the altitude of the sub-region, the greater the data missing rate, and the corresponding path decay rate is preset to a higher initial value.

After the decoder completes the economic growth prediction, the joint loss function integrates the prediction error, sparse regularization term, and prior consistency constraints of the causal direction of metadata, driving the end-to-end optimization of the entire model.

D. JOINT LOSS FUNCTION AND VARIATIONAL INFERENCE OPTIMIZATION

The joint loss function integrates economic growth prediction error, path sparsity regularization term, and prior consistency constraint of metadata causal direction, and uses variational inference to optimize the distribution of latent variables in the encoder output. Let the actual economic growth value be y, the decoder prediction value be $\hat{y}$, and the prediction error be expressed as mean squared error (MSE).

$$\mathcal{L}_{\text{pred}} = \frac{1}{N}\sum_{n=1}^{N}(y_n - \hat{y}_n)^2. \tag{8}$$

In formula (8), N is the total number of samples in the training set, y_n is the real economic growth indicator of the n-th sample, and $\hat{y}_n$ is the predicted value corresponding to the sample. The path sparsity regularization term $\mathcal{L}_{\text{sparse}}$ is defined as the L1 norm of the comprehensive importance coefficient λ_j of all candidate paths:

$$\mathcal{L}_{\text{sparse}} = \sum_{j=1}^{d_z} |\lambda_j| = \sum_{j=1}^{d_z} \|T_{:,j}\|_2 \cdot \rho_j^{-1}. \tag{9}$$

In formula (9), $T_{:,j}$ is the j-th column of the transition matrix, ρ_j^{-1} is the reciprocal of the path decay rate, and the L1 norm of their product compresses the weights of non-critical paths.

Metadata causal direction prior consistency constraint uses the pre-defined causal direction information in the metadata to impose a symbolic penalty on the path coefficients output by the decoder. Assuming that the metadata pre-specifies the expected causal symbol $s_j \in \{+1, -1, 0\}$ for each candidate path j, where $+1$ represents a positive impact, -1 represents a negative impact, and 0 represents no prior. The marginal effect symbol $\hat{s}_j = \mathrm{sign}(M_j)$ of path j on economic growth is extracted from the decoder, where M_j is the j-th column element of the output mapping matrix M. The consistency constraint loss is:

$$\mathcal{L}_{\text{prior}} = \sum_{j: s_j \neq 0} \max(0, -s_j \cdot M_j). \tag{10}$$

In formula (10), the penalty term is zero when s_j and M_j have the same sign, and positive when they have opposite signs, with its value proportional to the absolute value of M_j. This loss forces the path direction learned by the model to be consistent with the domain prior in the metadata.

The goal of variational inference is to maximize the lower log-marginal likelihood of the observed data, also known as the Evidence Lower Bound (ELBO). The encoder outputs an approximate posterior distribution $q_\phi(z \mid G)$ of the latent variable z, where ϕ is the encoder parameter, parameterized in diagonal Gaussian form as $N(\mu, \mathrm{diag}(\sigma^2))$. The prior distribution $p(z)$ is set to a standard Gaussian distribution $N(0, I)$. The ELBO includes a reconstruction term and a KL divergence term.

$$\begin{aligned}\mathcal{L}_{\text{ELBO}} &= \mathbb{E}_{q_\phi(z|G)}[\log p_\theta(y \mid z)] \\ &\quad - D_{\text{KL}}(q_\phi(z \mid G) \| p(z)).\end{aligned} \tag{11}$$

In formula (11), $p_\theta(y \mid z)$ is the conditional distribution defined by the decoder, whose mean is given by $\hat{y}$ and whose variance is fixed as a hyperparameter; $D_{\text{KL}}(\cdot\|\cdot)$ is the Kullback-Leibler divergence. The joint loss function $\mathcal{L}_{\text{total}}$

integrates the above terms and introduces the balancing hyperparameters η_{sparse} and η_{prior}:

$$\mathcal{L}_{\text{total}} = -\mathcal{L}_{\text{ELBO}} + \eta_{\text{sparse}}\mathcal{L}_{\text{sparse}} + \eta_{\text{prior}}\mathcal{L}_{\text{prior}}. \quad (12)$$

In formula (12), the negative sign transforms the ELBO maximization problem into a loss minimization problem. The optimization process employs a reparameterization technique to calculate the gradient: $z = \mu + \sigma \odot \epsilon$ is sampled from $q_\phi(z \mid G)$, where $\epsilon \sim N(0, I)$, making the sampling operation differentiable. The encoder parameters ϕ and decoder parameters θ of the entire model are jointly updated through stochastic gradient descent. In each iteration, a batch of heterogeneous graph samples is sampled from the training set, and z, $\hat{y}$, T, ρ, and M are calculated through forward propagation. Then, $\mathcal{L}_{\text{pred}}$, $\mathcal{L}_{\text{sparse}}$, $\mathcal{L}_{\text{prior}}$, and D_{KL} are calculated, and $\mathcal{L}_{\text{total}}$ is obtained by summing them. The parameters are then updated through backpropagation. After training, the path corresponding to the non-zero dimension in the latent variable z is the core transmission path from green finance to economic growth, and its coefficient is jointly given by M and T.

III. RESULTS AND DISCUSSION

A. EXPERIMENTAL ARRANGEMENT

The experiment is conducted using actual data from a specific region in a mountainous area inhabited by ethnic minorities (the border area of three prefectures on the eastern edge of the Hengduan Mountains). The dataset covers annual records from 2012 to 2022 and includes three sources: the first is green finance project data, including the annual funding amount, interest rate, and geographical coordinates of 78 green credit projects and 12 green bonds; the second is economic growth indicator data, including regional GDP and per capita net income with townships as the smallest statistical unit; the third is geographical and ethnic data, including elevation and slope raster data extracted from digital elevation model (DEM) and ethnic population census data. All data are standardized to the township scale, involving a total of 164 township samples. The metadata extraction process is as follows: for each green finance project, the project investment category, expected ecological benefits, and causal direction priors marked in the approval documents are extracted from its application documents; for geographical data, the spatial reference system and elevation accuracy are obtained from the metadata; for ethnic data, the ethnic composition category and language partition information of each township are extracted. When constructing the heterogeneous graph based on metadata, the total number of nodes is 164 geographic raster nodes, 78 green finance nodes, 12 industry sector aggregation nodes, 23 ethnic settlement nodes, and 11 time nodes (years), with 2416 edges.

The model hyperparameters are set as follows: number of graph convolutional layers $L = 3$, hidden state dimension $d_h = 128$ per layer; latent variable dimension $d_z = 48$; GRU time step $T_{\max} = 10$ years (covering the entire time window); initial learning rate of 0.001, decaying to 0.8 times the original value every 20 rounds; sparse regularization coefficient $\eta_{\text{sparse}} = 0.01$; prior consistency coefficient $\eta_{\text{prior}} = 0.05$; batch size of 8 heterogeneous graph subgraphs (each subgraph corresponds to a local graph structure of a township). The dataset is randomly divided into a training set (70%) (115 township samples), a validation set (15%) (24 township samples), and a test set (15%) (25 township samples), ensuring consistent altitude distribution across sub-regions during the partitioning process. The comparative models include: standard graphical convolutional network (GCN) direct regression, structural equation model (SEM) path analysis, and time-series fixed effects panel regression.

B. TRANSMISSION PATH COEFFICIENTS AND THEIR POSTERIOR UNCERTAINTY ANALYSIS

On a test set of 25 township samples, 2000 samples are drawn from the posterior distribution of the MD-DGVAE model to calculate the mean of the standardized path coefficients and its 95% highest density interval (HPD) for each candidate path. After sparsity regularization, 6 non-zero paths are retained from the initial 48-dimensional latent variables, corresponding to 6 core transmission paths from green finance to economic growth. The posterior mean of the path coefficient for the ecotourism infrastructure loan path (P1) is 0.324, with a 95% HPD range of [0.287, 0.361]; the mean coefficient for the clean energy small hydropower project path (P2) is 0.218, with a range of [0.179, 0.257]; the mean coefficient for the specialty organic agriculture credit path (P3) is 0.196, with a range of [0.163, 0.228]; the mean coefficient for the ecological compensation fund path (P4) is 0.152, with a range of [0.118, 0.185]; the mean coefficient for the green transportation subsidy path (P5) is 0.087, with a range of [0.051, 0.122]; and the mean coefficient for the ethnic handicraft green credit path (P6) is 0.071, with a range of [0.038, 0.103]. The 95% HPD ranges for all six paths do not include zero. The HPD interval width of P1 is 0.074, the widths of P3 and P6 are both 0.065, the width of P4 is 0.067, the width of P5 is 0.071, and the width of P2 is 0.078. Figure 2 shows the posterior mean of the path coefficients and the 95% HPD intervals for the six paths.

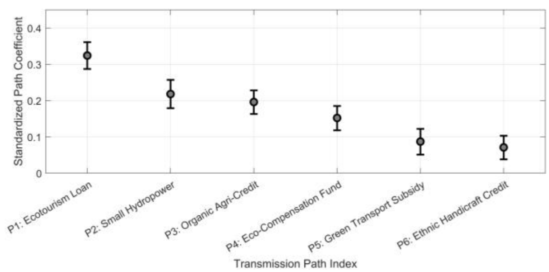

FIGURE 2. Posterior mean and 95% HPD interval of green finance transmission path coefficient.

C. ASSESSMENT OF THE ACCURACY OF ECONOMIC GROWTH FORECASTS

On a sample of 25 townships in the test set, the predicted economic growth value $\hat{y}$ output by the MD-DGVAE model is compared with the actual value y (measured by logarithmic GDP). The coefficient of determination R^2 is 0.8650, the mean absolute percentage error MAPE is 8.85%, and the root mean square error $\mathrm{RMSE}_{\log}$ is 0.0903. Figure 3 shows the scatter distribution of the actual and predicted values and the diagonal reference line on the test set.

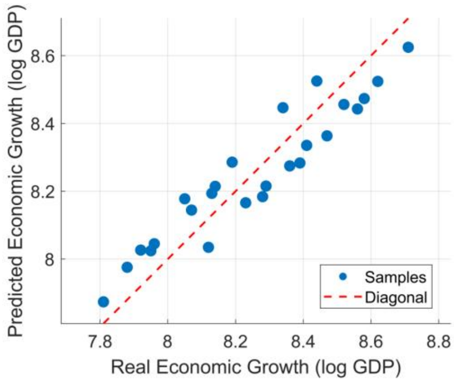

FIGURE 3. Scatter plot of actual and predicted economic growth values.

In Figure 3, each point represents a township sample in the test set, and the red dashed line is the diagonal line where the actual and predicted values are equal. All points are closely distributed on both sides of the diagonal, indicating that the model has high accuracy in predicting economic growth in the complex environment of ethnic minority mountainous areas. No systematic deviations are observed in the point cloud, verifying the ability of MD-DGVAE to effectively reconstruct economic growth indicators after capturing the transmission path of green finance.

D. PATH SPARSITY AND NETWORK STRUCTURE ANALYSIS

The MD-DGVAE model compresses the dimensionality of latent variables during training using L1 sparse regularization (coefficient $\eta_{\text{sparse}} = 0.01$). The initial encoder output latent variable dimension $d_z = 48$, meaning the number of candidate propagation paths is 48. After training, 42 of the 48 λ_j values are compressed to zero, while the combined importance coefficient λ_j of the remaining 6 paths remains non-zero, corresponding to the previously identified paths P1 to P6. The effective path retention rate is 12.5%. These 6 paths have a total of 124 edge connections in the initial heterogeneous graph, accounting for 5.13% of the total 2416 edges. The network density of the initial heterogeneous graph (288 nodes, 2416 edges) is 0.0585 (41328 edges in the complete graph). The subgraph containing only the 6 paths and their associated nodes (34 nodes in total) has 48 actual edges, and its network density is 0.086 (561 edges in the complete graph). Table 2 shows the time delay distribution, intermediate node types, and directional consistency of the six core transmission paths.

TABLE 2. Time Delay and Network Structure of Core Transmission Path

Path	Green finance direction	Economic growth indicator	Lag (yr)	Intermediary node types	Direction
P1	Ecotourism loan	Tourism revenue	2	Geo-grid → Industry sector	Positive
P2	Clean energy hydropower	Industrial added value	3	Geo-grid → Industry sector	Positive
P3	Organic agricultural credit	Primary industry output	2	Ethnic settlement → Geo-grid → Industry sector	Positive
P4	Ecological compensation fund	Transfer payment revenue	1	Geo-grid → Ethnic settlement	Positive
P5	Green transport subsidy	Transport output	4	Geo-grid → Industry sector	Positive
P6	Ethnic handicraft credit	Cultural industry output	3	Ethnic settlement → Industry sector	Positive

Table 2 shows that the time lags for the six paths range from 1 to 4 years, with P4 having the shortest lag (1 year) and P5 the longest (4 years). All paths are directional, consistent with the causal priors pre-defined in the metadata. Geographic raster nodes and ethnic settlement nodes appear as key intermediaries in each path, verifying the moderating effect of terrain complexity and ethnic distribution on the transmission of green finance. The six paths are independent of each other, with no overlap or merging observed.

IV. CONCLUSION

This paper addresses the challenge of quantifying the transmission pathways of green finance in ethnic minority mountainous areas by proposing a metadata-driven dynamic graph variational autoencoder model. The study reveals six core transmission pathways, including ecotourism loans and clean energy hydropower, and their time-lag characteristics, validating the crucial mediating roles of geographic grids and ethnic settlements in the transmission process. The model demonstrates robustness in economic growth prediction and pathway sparsity. Limitations include the lack of incorporation of nonlinear moderating effects from policy shifts and ethnic customs, and the reliance on expert annotation for metadata priors. Future research will introduce a dynamic metadata update mechanism and integrate ethnographic field survey data to enhance the model's adaptability and explanatory power in culturally diverse contexts.

FUNDING

This work was supported by Project of the Sichuan Research Center for Ethnic Mountainous Area Economic Development: Research on the Mechanism and Effect of Green Finance Empowering Economic Development in Sichuan Ethnic Mountainous Areas.

REFERENCES

[1] R. De Haas and A. Popov, "Finance and Green Growth," The Economic Journal, vol. 133, no. 650, pp. 637–668, 2023, doi: 10.1093/ej/ueac081.
[2] K. H. Wang, Y. X. Zhao, C. F. Jiang, and Z. Z. Li, "Does green finance inspire sustainable development? Evidence from a global perspective," Economic Analysis and Policy, vol. 75, pp. 412–426, 2022, doi: 10.1016/j.eap.2022.06.002.
[3] J. Gao, D. Wu, Q. Xiao, A. A. Randhawa, Q. Liu, and T. Zhang, "Green finance, environmental pollution and high-quality economic development—a study based on China's provincial panel data," Environmental Science and Pollution Research, vol. 30, no. 11, pp. 31954–31976, 2023, doi: 10.1007/s11356-022-24428-0.

[4] J. Wang, S. Zhang, Y. Xiao, and R. Song, "A Review on Graph Neural Network Methods in Financial Applications," Journal of Data Science, vol. 20, no. 2, pp. 111–134, 2022, doi: 10.6339/22-JDS1047.

[5] S. Bloemheuvel, J. van den Hoogen, D. Jozinović, A. Michelini, and M. Atzmueller, "Graph neural networks for multivariate time series regression with application to seismic data," International Journal of Data Science and Analytics, vol. 16, no. 3, pp. 317–332, 2023, doi: 10.1007/s41060-022-00349-6.

[6] M. Jin, H. Y. Koh, Q. Wen, D. Zambon, C. Alippi, G. I. Webb, I. King, and S. Pan, "A Survey on Graph Neural Networks for Time Series: Forecasting, Classification, Imputation, and Anomaly Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 12, pp. 10466–10485, 2024, doi: 10.1109/TPAMI.2024.3443141.

[7] Y. Wang, "Advances in spatiotemporal graph neural network prediction research," International Journal of Digital Earth, vol. 16, no. 1, pp. 2034–2066, 2023, doi: 10.1080/17538947.2023.2220610.

[8] Z. Ye, Y. J. Kumar, G. O. Sing, F. Song, and J. Wang, "A Comprehensive Survey of Graph Neural Networks for Knowledge Graphs," IEEE Access, vol. 10, no. 1, pp. 75729–75741, 2022, doi: 10.1109/ACCESS.2022.3191784.

[9] M. Masmoudi, S. Ben Abdallah Ben Lamine, M. H. Karray, et al., "Semantic data integration and querying: A survey and challenges," ACM Computing Surveys, vol. 56, no. 8, pp. 1–35, 2024, doi: 10.1145/3653317.

[10] L. Zhong, J. Wu, Q. Li, et al., "A comprehensive survey on automatic knowledge graph construction," ACM Computing Surveys, vol. 56, no. 4, pp. 1–62, 2023, doi: 10.1145/3618295.

[11] M. Hofer, D. Obraczka, A. Saeedi, H. Köpcke, and E. Rahm, "Construction of Knowledge Graphs: Current State and Challenges," Information, vol. 15, no. 8, pp. 509–517, 2024, doi: 10.3390/info15080509.

[12] R. Bing, G. Yuan, M. Zhu, F. Meng, H. Ma, and S. Qiao, "Heterogeneous graph neural networks analysis: a survey of techniques, evaluations and applications," Artificial Intelligence Review, vol. 56, no. 8, pp. 8003–8042, 2023, doi: 10.1007/s10462-022-10375-2.

[13] N. Mrabah, M. Bouguessa, M. F. Touati, and R. Ksantini, "Rethinking Graph Auto-Encoder Models for Attributed Graph Clustering," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 9, pp. 9037–9053, 2023, doi: 10.1109/TKDE.2022.3220948.