Supplement of Earth Syst. Sci. Data, 17, 4005–4022, 2025
https://doi.org/10.5194/essd-17-4005-2025-supplement

Open Access Earth System Science Data

Supplement of

A vegetation phenology dataset developed by integrating multiple sources using the reliability ensemble averaging method

Yishuo Cui et al.

Correspondence to: Yuyu Zhou (yuyuzhou@hku.hk) and Yongshuo H. Fu (yfu@bnu.edu.cn)

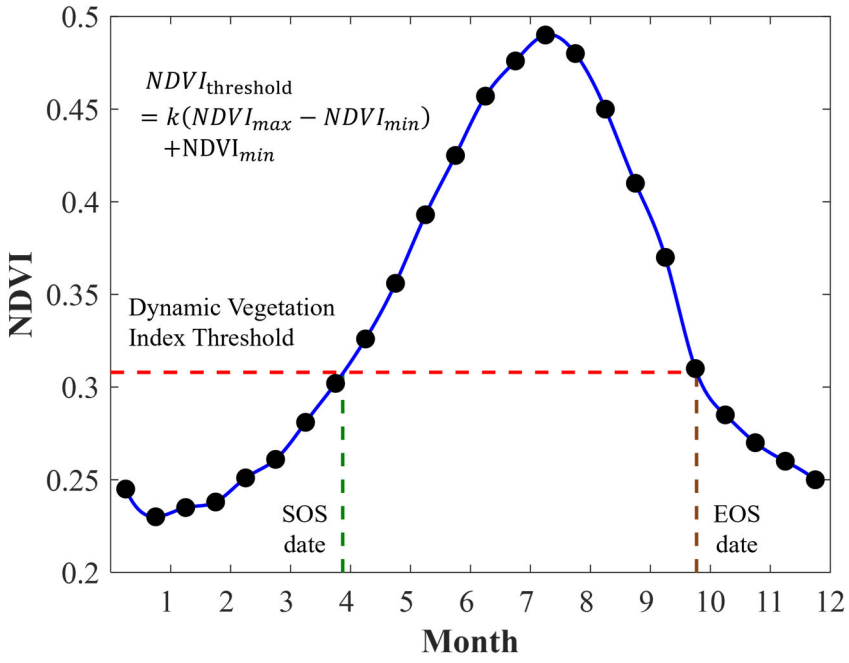

Figure S1: Diagram of vegatation phenology extraction using the threshold method. The red dotted line denotes the dynamic vegetation index threshold, which is calculated by multiplying the given threshold by the range of vegetation index then plus the minimum of the vegetation index. The green dotted line denotes the SOS (start of season) date, and the brown dotted line denotes the EOS (end of season) date.

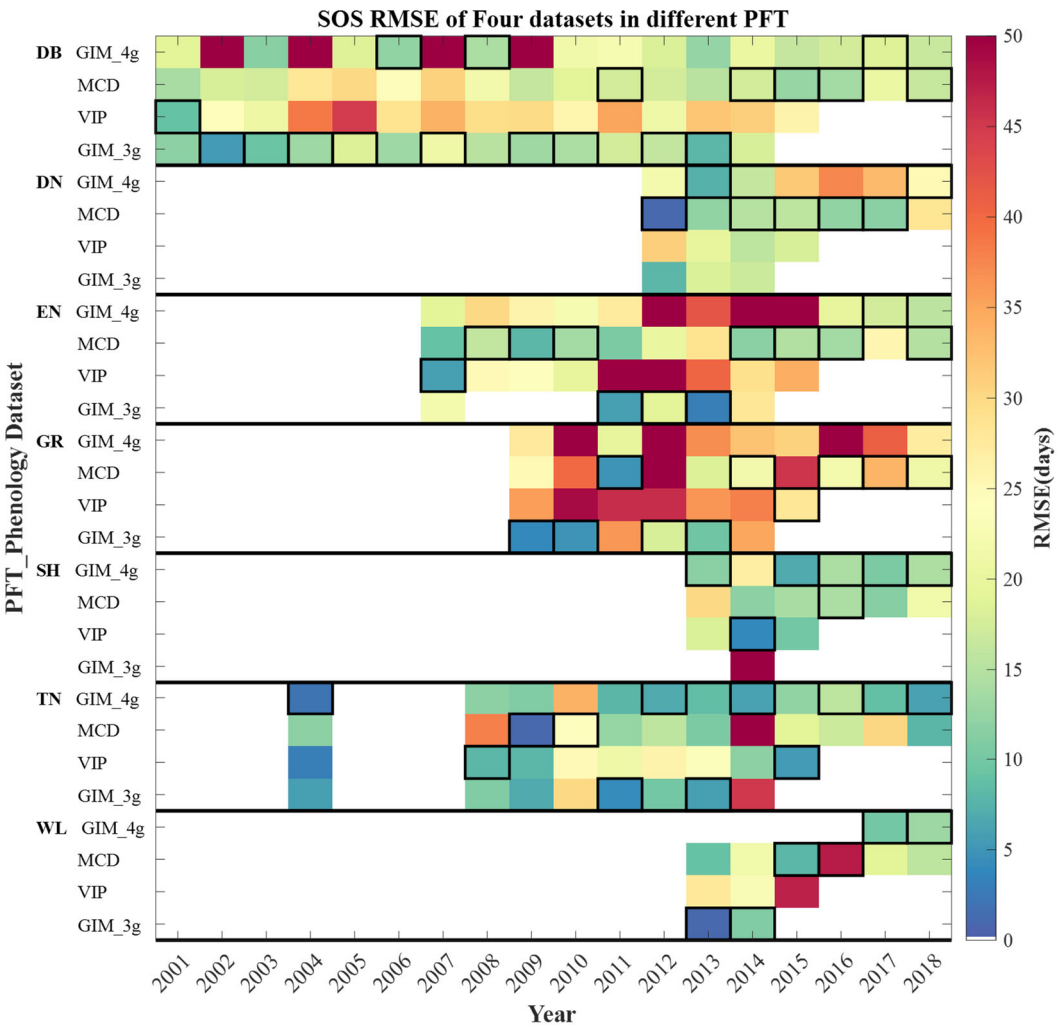

Figure S2: RMSE between the four remote sensing datasets and the PhenoCam dataset for SOS across different PFTs during 2001 – 2018. Four datasets refer to GIM_4g, MCD12Q2, VIP, and GIM_3g datasets, respectively. PFT: plant functional type, DB: deciduous broadleaf, DN: deciduous needleleaf, EN: evergreen needleleaf, GR: grassland, SH: shrubs, TN: tundra, WT: wetland. The black boxes represent the best data for the year in that PFT.

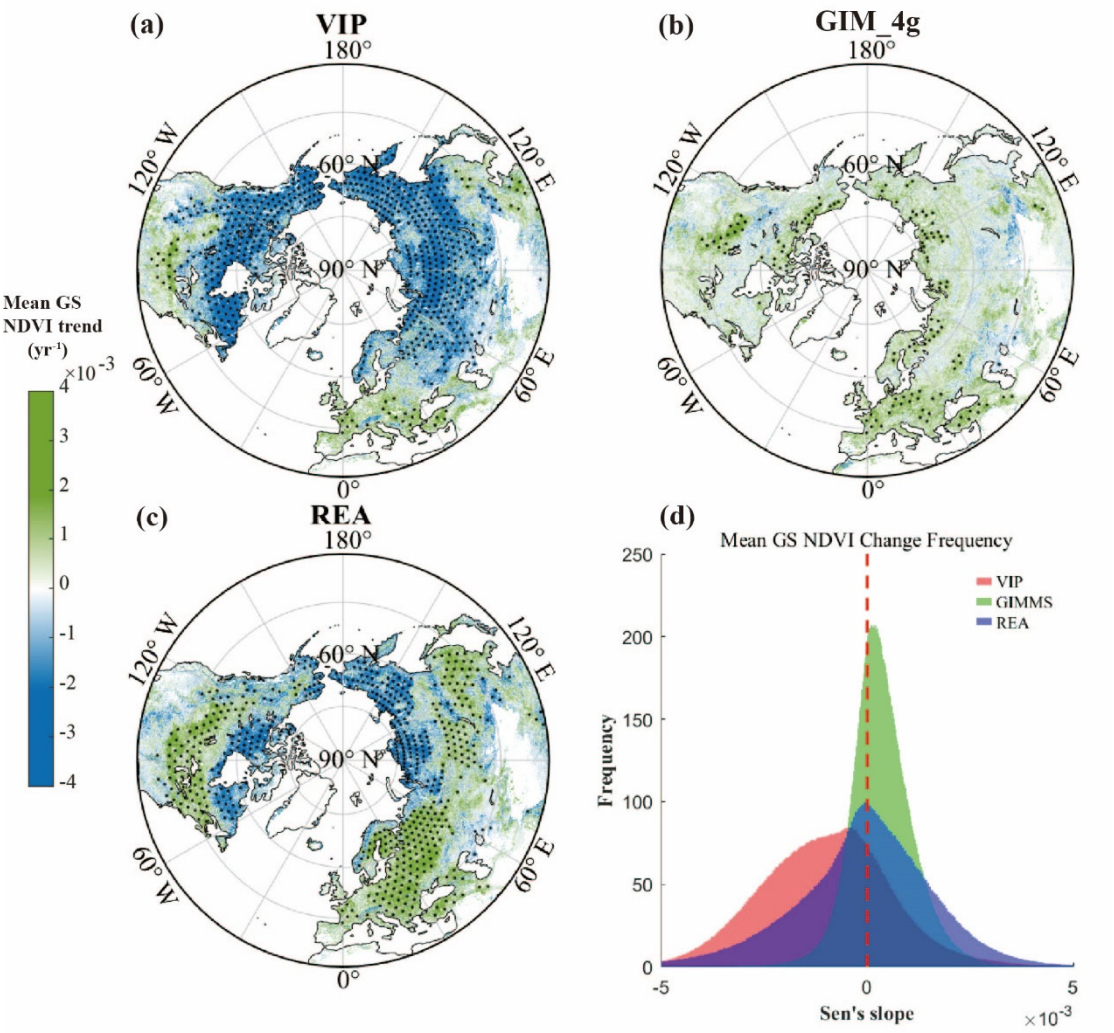

Figure S3: Trends in growing season NDVI derived from (a) the VIP phenology dataset, (b) the GIM_4g dataset, and (c) the REA phenology dataset for the period 1982 – 2015. Panel (d) presents the frequency distribution of the Sen's slope for the growing season NDVI trend in regions north of 30°N. A positive Sen's slope indicates an increasing (greening) trend, while a negative slope indicates a decreasing (browning) trend in seasonal vegetation greenness. GS: growing season. Black dots indicate regions where the trend is statistically significant ($P < 0.05$).

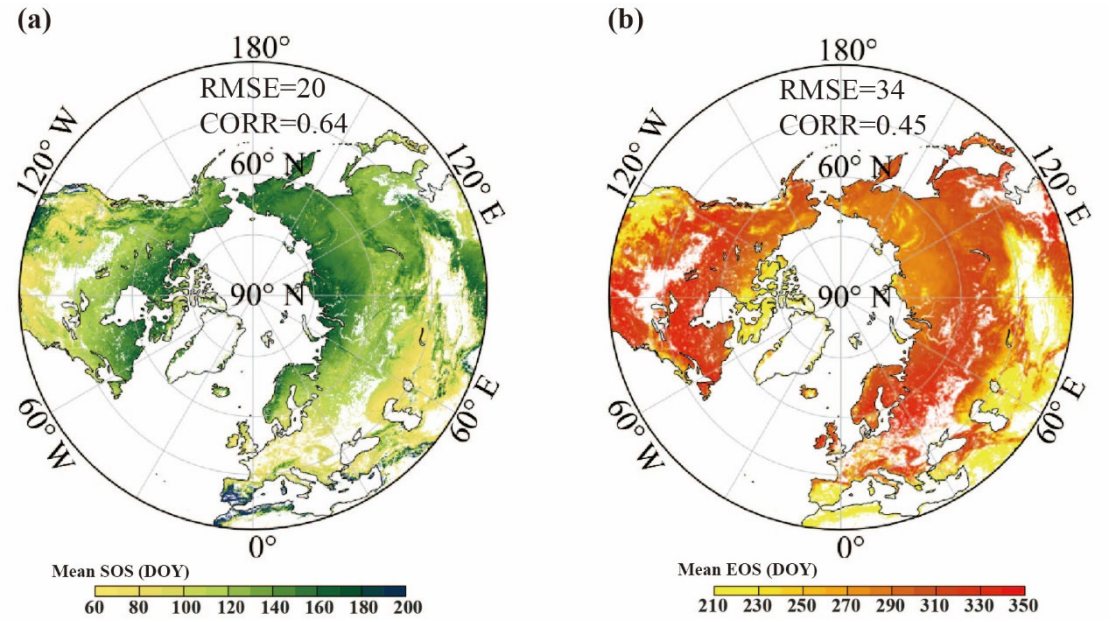

Figure S4: Mean (a) SOS and (b) EOS dates (DOY) obtained using the simple average across all four datasets for the period 1982 – 2020. RMSE represents the root mean square error (days), and CORR represents the correlation coefficient. The four datasets include MCD12Q2, VIP, GIM_4g, and GIM_3g. For further details on these datasets, please refer to Section 2.1.

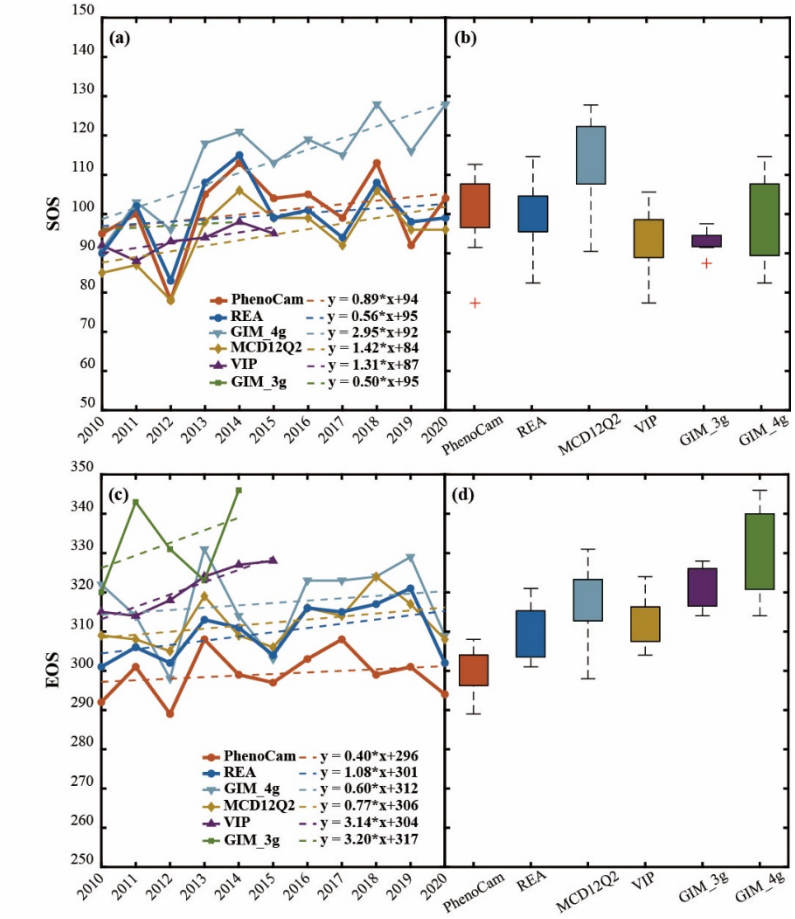

Figure S5: Time series and box plots of Morganmonroe site data with each phenology dataset and the merged phenology dataset obtained using the REA method. (a-b) SOS time series and box plot of the PhenoCam, GIM_4g,

MCD12Q2, VIP, GIM_3g, and REA datasets, respectively, (c-d) EOS time series and box plot of the PhenoCam, GIM_4g, MCD12Q2, VIP, GIM_3g, and REA datasets, respectively. The dotted lines in a and c represent the trendlines for each dataset. The Morganmonroe site, located in Morgan Monroe State Forest, Indiana, is characterized by deciduous broadleaf forest.

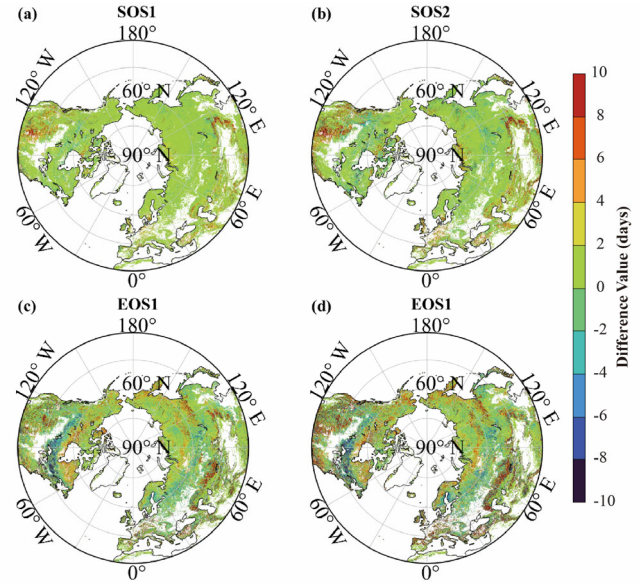

Figure S6: Comparison of fusion results of four groups of data at different time lengths from 2001 to 2010. (a, c) Difference plots between the fusion results of SOS and EOS for the four datasets over the 2001 – 2010 period and the original long-term REA fusion results (1982 – 2020). (b, d) Difference plots between the fusion results of SOS and EOS obtained by separately fusing the two sub-periods of data (2001 – 2005 and 2006 – 2010) and subsequently concatenating them, and the original long-term REA fusion results.

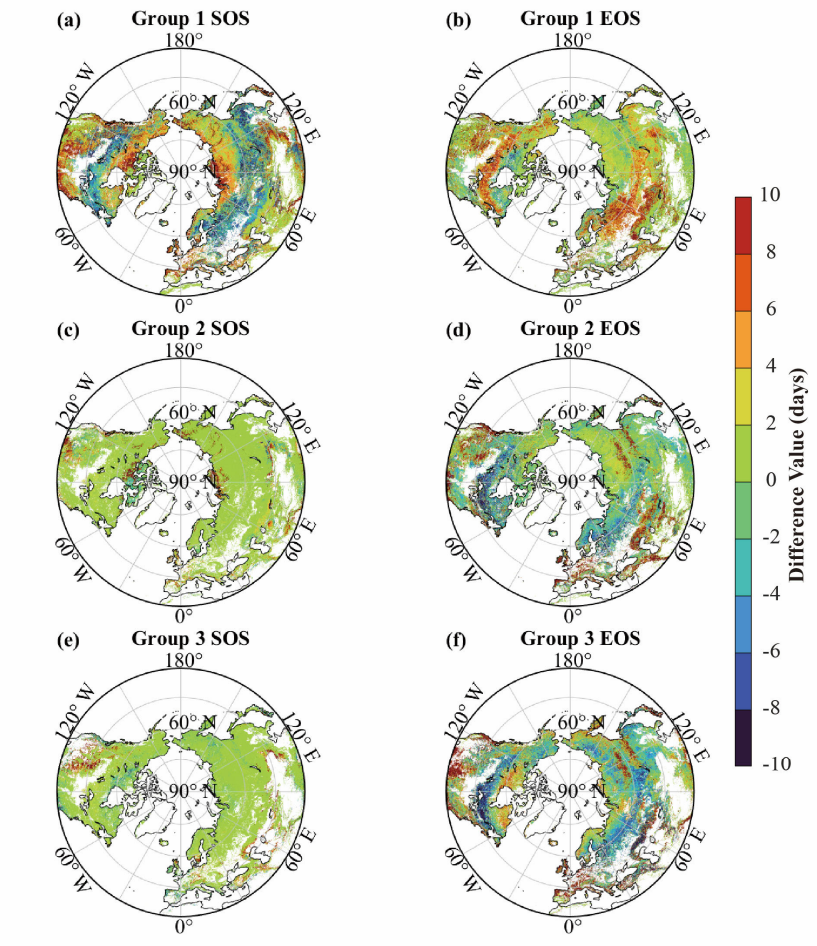

Figure S7: Differences between the SOS and EOS data fusion results from various data source combinations and the fusion results of the long-term four datasets over the 2001-2010 period. (a-b) Group 1 represents the difference between the fusion of GIM_4g and VIP data and four datasets REA result. (c-d) Group 2 represents the difference between the fusion of VIP data and GIM_3g data and four datasets REA result. (e-f) Group 3 represents the difference between the fusion of GIM_3g and MCD12Q2 data and four datasets REA result.

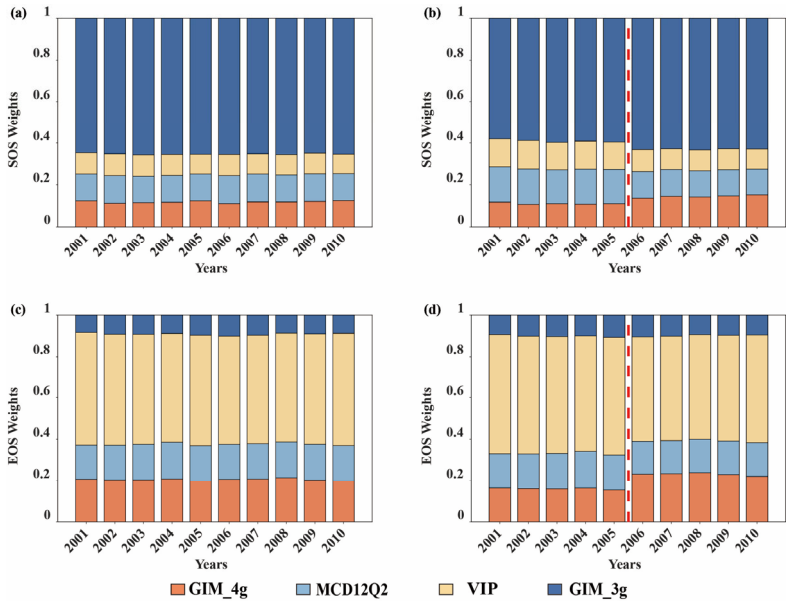

Figure S8: Comparison of fusion data weights for the four datasets over different time lengths from 2001 to 2010. (a, c) Weights of the SOS and EOS fusion data for the four datasets over the entire 2001 – 2010 period. (b, d) Weights of the fusion data after separately fusing the two sub-periods of data (2001-2005 and 2006-2010) and subsequently concatenating them. The red dashed line separates the two sub-periods.

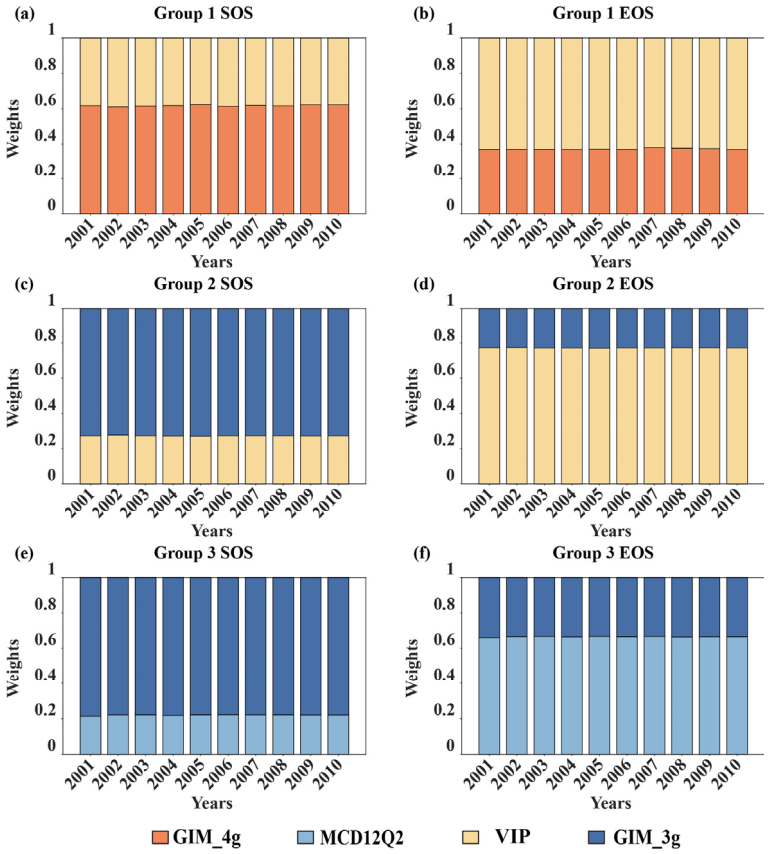

Figure S9: Distribution of fusion weights for SOS and EOS data from different data source combinations over the 2001 – 2010 period. (a, b) Group 1: Fusion weights for the combination of GIM_4g and VIP data. (c, d) Group 2: Fusion weights for the combination of VIP data and GIM_3g data. (e, f) Group 3: Fusion weights for the combination of GIM_3g and MCD12Q2 data.