

Modeling the spatial distribution of carbon dioxide concentration by machine learning

Sainbayar Dalantai^{1,*}, Erdenesukh Sumiya², Amarsaikhan Damdinsuren¹

¹*Division of Aerospace Remote Sensing, Institute of Geography and Geoecology, Mongolian Academy of Sciences, Ulaanbaatar, Mongolia*

²*Department of Meteorology and Hydrology, National University of Mongolia, Ulaanbaatar, Mongolia*

*Corresponding author email and ORCID: sainbayard@mas.ac.mn, <https://orcid.org/0000-0001-8806-6167>

Received: 09 Jul 2025 / Accepted: 17 Oct 2025 / Published online: 19 December 2025

ABSTRACT

The rapid increase in atmospheric carbon dioxide (CO₂) is a significant contributor to modern climate change and global warming. High-resolution data are essential for scientifically valid studies of CO₂ dynamics and changes. Although greenhouse gas observation satellites exist, they are still inadequate in terms of spatiotemporal resolution. We modeled to obtain a high-spatiotemporal resolution (16-day, 0.1x0.1) XCO₂ by machine learning algorithms (such as Random Forest (RF), eXtreme Gradient Boosting (XGB), Light Gradient Boosting Machine (LGBM), and Stacking (Stk)) using multisource data (such as weather parameters, vegetation index, land cover, elevation and population density) over Mongolia and China from 2010 to 2022. The results indicated that the XGB demonstrated good performance, with R² values exceeding 0.95 and 10-fold cross-validation R² (cv) values exceeding 0.55 during the study periods. Additionally, when calculating the Shapley Additive Explanations (SHAP) values based on the XGB model's results, the weather variables (53%) had the largest influence on the distribution of XCO₂. In comparison, population density alone had an 11% influence. As a result, the average XGB XCO₂ over the study period was 402.3±1.14 ppm, with the highest value observed in eastern China and the lowest value observed in Mongolia and the Qinghai-Tibet Plateau. This is related to the region's economic development and human activities. In terms of seasonal dynamics, the XGB XCO₂ values for winter, spring, summer, and autumn were 403.4±1.2 ppm, 403.3±1.3 ppm, 400.9±1.5 ppm, and 402.1±1.0 ppm, respectively. This shows that XCO₂ is highest in winter and lowest in summer. This seasonal dynamic is because plants absorb more CO₂ through photosynthesis in summer than in other seasons. The atmospheric XCO₂ map produced in this study can serve as a baseline map for studying the carbon cycle in terrestrial ecosystems and climate change.

Keywords: Machine learning, Carbon dioxide, Climate change, Greenhouse gas, Shapley Additive Explanations (SHAP)

Нүүрсхүчлийн хийн агууламжийн орон зайн тархалтыг машин сургалтаар загварчлах нь

Сайнбаяр Далантай^{1,*}, Эрдэнэсүх Сумъяа², Амарсайхан Дамдинсүрэн¹

¹Агаар, сансрын зайнаас тандан судлалын салбар, Газарзүй, геоэкологийн хүрээлэн, Шинжлэх Ухааны Академи, Улаанбаатар, Монгол

²Цаг уур судлалын тэнхим, Монгол Улсын Их Сургууль, Улаанбаатар, Монгол

*Холбоо барих зохиогчийн цахим хаяг, ORCID: sainbayard@mas.ac.mn, <https://orcid.org/0000-0001-8806-6167>

Хүлээн авсан: 2025 оны 07 сарын 09 өдөр / Зөвшөөрөгдсөн: 2025 оны 10 сарын 17 өдөр / Нийтлэгдсэн: 2025 оны 12 сарын 19 өдөр

ХУРААНГУЙ

Агаар мандал дахь нүүрсхүчлийн хий (CO_2)-н агууламжийн хурдацтай өсөлт нь Дэлхийн дулаарал, орчин үеийн уур амьсгалын өөрчлөлтийн үндсэн шалтгаан болж байна. CO_2 -н динамик, өөрчлөлтийн талаарх өндөр нарийвчлалтай өгөгдөлд тулгуурласан судалгаа хийх нь чухал юм. Хэдийгээр хүлэмжийн хийн ажиглалтын хиймэл дагуулууд байгаа боловч цаг хугацаа, орон зайн хувьд шаардлага хангахгүй хэвээр байна. Иймд энэхүү судалгаандаа, бид Монгол болон БНХАУ (Бүгд Найрамдах Хятад Ард Улс)-ын нутаг дэвсгэрийн хэмжээнд, 2010-аас 2022 он хүртэлх хугацааны Хүлэмжийн хийн ажиглалтын хиймэл дагуул (GOSAT)-ын хуурай агаарын босоо баганын дундаж XCO_2 болон цаг уурын хэмжигдэхүүнүүд, ургамлын индекс, газрын бүрхэвч, гадаргын өндөршил, хүн амын нягтшил зэрэг 12 өгөгдлийг ашиглан машин сургалтын 4 өөр аргаар (Random Forest (RF), eXtreme Gradient Boosting (XGB), Light Gradient Boosting Machine (LGBM), Stacking (Stk)) XCO_2 -г загварчлан харьцуулсан. Эдгээрээс XGB XCO_2 загвар хамгийн өндөр гүйцэтгэлтэй буюу детерминацийн коэффициент $R^2 > 0.95$, солбин бататгах $R^2(\text{cv}) > 0.55$ байлаа. Түүнчлэн XGB загварын үр дүнд тулгуурлан Таамаглалыг тайлбарлах нэгдсэн тогтолцоо буюу Shapley Additive Explanations (SHAP) утгыг тооцоход XCO_2 -н тархалтад цаг уурын хувьсагчид хамгийн их нөлөө (53%) үзүүлсэн бол хүн амын нягтшил дангаараа 11% нөлөө үзүүлсэн. Өндөр гүйцэтгэлтэй XGB загвараар судалгааны талбайг хамарсан өндөр нарийвчлалтай (16° өдрийн, $0.1^\circ \times 0.1^\circ$) XGB XCO_2 -н орон зайн тасралтгүй өгөгдлийг гарган авсан. Үр дүнд судалгааны хугацааны талбайн дундаж XGB XCO_2 нь 402.3 ± 1.14 ppm, их утга нь БНХАУ-ын зүүн хэсгээр, бага утга нь Монгол болон Чинхай-Төвөдийн өндөрлөгт ажиглагдсан. Энэ нь бүс нутгийн эдийн засгийн хөгжил, хүний үйл ажиллагаатай холбоотой юм. Улирлын динамикийн хувьд XGB XCO_2 утга нь өвөл, хавар, зун, намрын улиралд харгалзан 403.4 ± 1.2 ppm, 403.3 ± 1.3 ppm, 400.9 ± 1.5 ppm, 402.1 ± 1.0 ppm байлаа. Эндээс үзвэл XCO_2 нь өвлийн улиралд хамгийн их утгатай, зун хамгийн бага утгатай байна. Улирлын динамик нь зуны улиралд ургамал фотосинтезийн процессоор CO_2 -г бусад улиралтай харьцуулахад ихээр шингээдэгтэй холбоотой. Энэхүү судалгааны ажлын үр дүн нь биосферийн нүүрстөрөгчийн эргэлтийг загварчлах, уур амьсгалын өөрчлөлтийг судлахад суурь өгөгдөл болно.

Түлхүүр үгс: Машин сургалт, Нүүрсхүчлийн хий, Уур амьсгалын өөрчлөлт, Хүлэмжийн хий, Таамаглалыг тайлбарлах нэгдсэн тогтолцоо (SHAP)

1. ОРШИЛ

Уур амьсгалын өөрчлөлтийн үндсэн шалтгаан болсон агаар мандал дахь CO_2 -н дундаж 1750 онд 278 ppm байсан бол 2021 онд 415 ppm болж нэмэгдсэн нь сүүлийн 2 сая жилд ажиглагдаагүй өндөр түвшинд хүрсэн үзүүлэлт [1] бөгөөд энэ нь сүүлийн 800000 жилд тохиолдсон өсөлтөөс даруй 10 дахин их хурдацтайгаар өссөн үзүүлэлт юм [2]. Агаарын мандлын хийн найрлагын өөрчлөлт, ялангуяа хүлэмжийн хийн өөрчлөлтийн судалгаа бол хүрээлэн буй орчны физик, хими, цаг уур, геологи, далай судлал болон тоон загварчлал, системийн загварчлал зэрэг олон салбарын шинжлэх ухааныг хамарч, Дэлхийн системд явагдах биогеохимийн хүрээнд судлах шаардлагатай судлагдахуун болоод байна [3]. Хэдийгээр нүүрсхүчлийн хий, метан зэрэг хийнүүдийн мониторинг болон хиймэл дагуулын ажиглалтын мэдээ байгаа хэдий ч цаг хугацаа, орон зайн тархалтын хувьд хэрэгцээ, шаардлагыг бүрэн хангахгүй байна [4].

Тухайлбал, CO_2 -г хэмжих анхны зорилтод хиймэл дагуулууд болох GOSAT-г 2009 онд, OCO-2 (Orbiting Carbon Observatory-2) хиймэл дагуулыг 2014 онд тус тус хөөргөсөн байдаг [5], [6]. Үүний дараагаар цувралаар GOSAT-2 хиймэл дагуулыг 2018 онд, OCO-3-г 2019 онд хөөргөсөн. Эдгээр хиймэл дагуулууд нь бусад хиймэл дагуултай харьцуулахад газрын гадарга орчмын нүүрсхүчлийн хийн тархалт, ялгарлын эх үүсвэрийг сайн илэрхийлдэг [7]. Газрын станцын мэдээтэй харьцуулахад хэмжилтийн нарийвчлалын хувьд OCO хиймэл дагуулын цувралууд нь GOSAT хиймэл дагуулын цувралуудтай харьцуулахад алдаа багатайгаар хэмждэг байна (дэлгэрэнгүйг [5]-аас харна уу). Харин GOSAT-1, OCO-2 нь дараа үеийн цувралтайгаа харьцуулахад хэмжилтийн алдаа бага, харьцангуй өндөр нарийвчлалтайгаар зайнаас танддаг болохыг судлаачид тогтоосон. Энэ нь хиймэл дагуулуудын сенсор болон алгоритмууд нь илүү тогтвортой байж болох юм гэжээ [5].

Түүнчлэн уур амьсгалын системийн загваруудад хүрээлэн буй орчны тогтворгүй байдал, хүний үйл ажиллагаа зэргээс шалтгаалан түүхэн өгөгдлийн динамик өөрчлөлт, учир шалтгааны холбоосны таамаглал зөрчилдөх тохиолдол гардаг зэрэг нь загварын тодорхойгүй байдлыг ихэсгэх шалтгаан болдог [8]. Иймд CO_2 -н хэлбэлзлийг судлахдаа аль болох олон хүчин зүйлийг (системүүдийг төлөөлөх үзүүлэлтүүд) авч, хиймэл оюун, машин сургалтын алгоритмаар загварчлах нь үр дүнтэй шийдэл байх болно.

Сүүлийн жилүүдэд судлаачид зайнаас тандсан XCO_2 өгөгдлийг олон эх сурвалжийн өгөгдөлтэй хослуулан машин сургалтын алгоритмаар загварчилж өндөр нарийвчлалтай өгөгдөл гарган авах судалгааны ажлууд хийх болсон. Тухайлбал, He ба бусад (2023) OCO-2 XCO_2 өгөгдлийг ашиглан БНХАУ-ын хэмжээнд машин сургалтын ансамбль алгоритмууд (RF, extreme random forest-ERF, XGB, LGBM)-аар загварчлахад RF хамгийн өндөр гүйцэтгэлтэй байсан. Загварт үл хамаарах хувьсагчдын ач холбогдлыг үнэлэхэд СТ (CarbonTracker) XCO_2 хамгийн чухал нөлөөлөгч буюу харьцангуй ач холбогдол нь 83% байсан бол цаг уурын хэмжигдэхүүнүүд 9%, хүн амын нягтшил, ургамалжилт болон газар ашиглалт 2%-ийн нөлөөг үзүүлсэн болохыг тооцжээ [9]. He ба бусад (2022) адил өгөгдлийг ашиглан машин сургалтын LGBM алгоритмаар загварчилж, газрын станцуудын мэдээгээр шалгасан. Тус загварт цаг уурын хэмжигдэхүүнүүд 16%, газар ашиглалт 15% нөлөө үзүүлсэн бол СТ XCO_2 хамгийн их буюу 44% нөлөө үзүүлжээ [10]. Түүнчлэн Дэлхийн хэмжээнд Li ба бусад (2022) машин сургалтын ETR алгоритмыг ашигласан байна. Тэрээр XCO_2 -д нөлөөлөх хүчин зүйлээс LAI, NPP зэрэг нь ургамлын амьсгал, фотосинтезийн процесстэй холбоотойгоор CO_2 урсгалыг тодорхойлоход онцгой ач холбогдолтой төдийгүй, Дэлхийн бөмбөрцгийн хойд хэсэгт улирлын явц нь ургамлын ургах үеийн динамикаас хамаарч байгааг дурджээ [11]. Монгол улсын нутаг дэвсгэрийг хамарсан GOSAT XCO_2 -н судалгаа нь геостатистик арга дээр тулгуурласан байна [12], [13], [14]. Геостатистик интерполяцийн арга нь орон зай, цаг хугацааны шийд нь ижил байх олон оролтын өгөгдлийг шаардлагаас гаралтын үр дүн нь хугацааны шийдээр хязгаарлагддаг төдийгүй орон зайн хувьд жигдрүүлэх (smooth) сул талтай.

Эдгээр өмнөх судалгаанууд нь ихэвчлэн машин сургалтын нэгэн төрлийн загварууд дээр тулгуурласнаас гадна, OCO-2 XCO_2 өгөгдлийг ашиглажээ. OCO-2 XCO_2 өгөгдлийн орон зайн шийд ($1.29 \times 2.25\text{км}$) нь GOSAT XCO_2 өгөгдөлтэй харьцуулахад өндөр боловч хугацааны шийд багатайгаас гадна өгөгдлийн цуваа нь богино (2014 оноос хойш) юм. Иймд бид судалгаандаа 2010-2022 он хүртэлх хугацааны 13 жилийн GOSAT XCO_2 ашиглан машин сургалтын 4 өөр алгоритмаар загварчлан харьцуулж, эдгээрээс өндөр гүйцэтгэлтэй загварт тулгуурлан орон зай, цаг хугацааны өндөр нарийвчлалтай (16 өдөр, 0.1×0.1) XCO_2 -н орон зайн тасралтгүй зургийг гарган авахыг зорилоо. Үүнээс гадна бид судалгаандаа SHAP утгыг тооцож, үл хамаарах хувьсагчдын загварт оруулсан хувь нэмрийг

үнэлж, цаашлаад хүчин зүйл хоорондын харилцан шүтэлцээ, физик процессыг тайлбарлана.

2. СУДАЛГААНЫ ТАЛБАЙ, АШИГЛАСАН ӨГӨГДӨЛ

2.1. Судалгааны талбай

Судалгааны талбай нь хойд өргөргийн 18°9'-53°34', зүүн уртрагийн 73°33'-134°46'-ийн хооронд Зүүн Азийн 10.9 сая.км² нутаг дэвсгэрийг хамарна. Тус талбайд Монгол улс болон БНХАУ (энд арлуудыг оруулаагүй) оршдог. 190 гаруй орныг хамарсан нэг хүнд ногдох хүлэмжийн хийн ялгарлын хандлагын судалгаагаар хамгийн их өсөх хандлагатай 20 улс нь Ази тивд оршдог төдийгүй, энэ нь хүн амын төвлөрөл, эдийн засгийн өсөлттэй холбоотойг дурдсан байдаг [15]. Тус судалгаагаар 2004 оноос хойш нэг хүнд ногдох хүлэмжийн хийн ялгаралд хамгийн их нөлөөлсөн нь Ази тивээс Монгол улс, БНХАУ болохыг тогтоожээ. Иймд бидний судалгааны талбай нь Дэлхийн хүлэмжийн хийн ялгарал болон нэг хүнд ногдох ялгарлын хувьд ихээхэн анхаарал татаж байгаа юм.

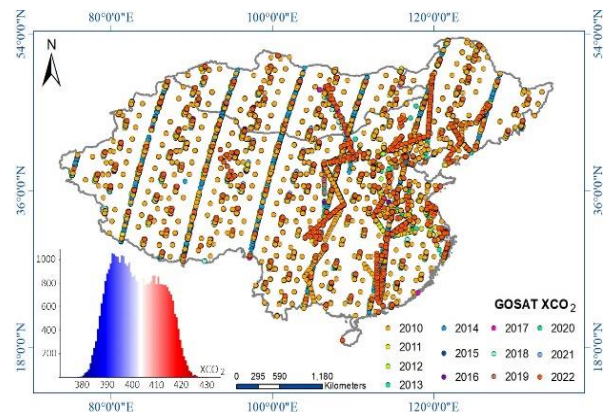
2.2. GOSAT XCO₂-н өгөгдөл

GOSAT нь агаар мандал дахь XCO₂ болон XCH₄ ажиглалтын анхны хиймэл дагуул бөгөөд 2009 оны 1-р сарын 23-нд Японы сансар судлалын агентлаг (JAXA)-аас хөөргөсөн [16]. Тус хиймэл дагуулын хугацааны шийд нь 3 өдөр, орон зайн шийд нь 10.5 км диаметр юм [17]. Судалгаандаа GOSAT FTS SWIR Level2 бүтээгдэхүүний засварлагдсан 2010-2022 оны 104898 мөр мэдээлэл бүхий өгөгдлийг <https://data2.gosat.nies.go.jp> сайтаас татаж авсан. Эдгээр өгөгдлийг ArcGIS программ хангамжийг ашиглан 0.1x0.1 торлол үүсгэж, торлолын нэг нүдэнд хамрагдах мэдээллийг 16 өдрөөр дундажлан 38891 мөр мэдээлэл бүхий XCO₂ үл хамаарах хувьсагчийг гарган авсан. **Зураг 1**-д судалгаанд ашигласан GOSAT XCO₂ өгөгдлийн орон зайн болон утгын гистограммыг харуулав. XCO₂ утгын гистограммаас үзэхэд 402 ppm-ээс бага утга нийт өгөгдлийн цувааны 49.3%-ийг эзэлсэн. Статистик үр дүнг авч үзвэл судалгааны хугацаанд хамгийн их, бага, дундаж ±стандарт хазайлт нь харгалзан, 433.18 ppm, 373.99 ppm, 401.89±9.79 ppm байна.

2.3. Бусад өгөгдөл

Агаар мандал дахь хийн хольцын харьцааны динамик хэлбэлзэл, өөрчлөлт нь олон хүчин зүйлээс (физик, хими, биогеохимийн процесс,

хүний үйл ажиллагаа зэрэг) хамаарсан хам процесс юм [18]. Агаар дахь CO₂-н агууламжийн орон зай, цаг хугацааны тархалт, динамик хэлбэлзэлд ургамал бүрхэвч, цаг уурын хэмжигдэхүүнүүд болон хүний үйл ажиллагаатай холбоотой хүчин зүйлүүд гэсэн 3 үндсэн категорид хуваагдах хувьсагчид шууд болон хам байдлаар нөлөөлдөг болохыг цөөнгүй судлаачид дурдсан байдаг [18], [19]. Бид судалгаандаа ургамлын индекс, цаг уурын хэмжигдэхүүн, хүн амын нягтшил болон бусад гэсэн категорид хамаарах нийт 12 үл хамаарах хувьсагчдыг судалгааны хамрах хугацаагаар авсан. Судалгаанд ашигласан өгөгдлийн мэдээллийг **Хүснэгт 1**-д харууллаа.



Зураг 1. GOSAT XCO₂ өгөгдөл

Хүснэгт 1. Судалгаанд ашигласан өгөгдлийн мэдээлэл

Хувьсагчийн нэр (товчлол)	Орон зай/ хугацааны шийд
Ургамлын нормчлогдсон ялгарлын индекс (NDVI)	1 км x 1 км/16 өдөр
Сайжруулсан ургамлын индекс (EVI)	
Гадарга орчмын агаарын даралт (SP)	0.1° x 0.1°/цаг
Хөрсөн дэх усны агууламж, 0-7 см (SWC)	
Хязгаарын үе давхаргын өндөр (BLH)	0.25° x .25°/цаг
Нарны цацрагийн эрчимшил (Rad)	
Агаарын температур (T2m)	0.1°x0.1°/өдөр
Шүүдэр цэгийн температур (Dewt)	
10 м-ийн салхины хурд (WS)	1 км x 1 км/жил
Хүн амын нягтшил (Pop)	
Газрын бүрхэвч (LC)	500 м x 500 м/жил
Гадаргын өндөршил (DEM)	0.03°x0.03°

Судалгаанд MODIS хиймэл дагуулын ургамлын индексийн MOD13A2 бүтээгдэхүүний

NDVI, EVI болон газрын бүрхэвчийн ангиллын MCD12Q1 бүтээгдэхүүний ургамлын функциональ бүлгүүд (PFT) дээр тулгуурласан ангиллын өгөгдлийг ашигласан. Цаг уурын хэмжигдэхүүний хувьд судалгаандаа Европын дунд хугацааны прогнозын төв (ECMWF)-ийн реанализийн (ERA-5) өгөгдлийн сангаас загварт орон зай, цаг хугацааны шийд нь хамгийн тохиромжтой байх нөхцөлөөр дахин загварчилсан өгөгдлүүдийг ашигласан (**Хүснэгт 1**). Үүнд, ERA-5 өгөгдлийн орон зайн шийдийг сайжруулсан ERA5-Land өгөгдлийн сангаас SP, SWC [20], хөдөө аж ахуй, экологийн судалгаанд зориулан дахин загварчилсан AgERA5 өгөгдлийн сангаас Rad, T2m, Dewt, WS өгөгдлүүд юм [21]. Түүнчлэн BLN өгөгдлийг ERA-5 өгөгдлийн сангаас [22] авч ашигласан. Үүнээс гадна судалгаанд Америкийн Нэгдсэн Улсын Агаар мандал, далай судлалын үндэсний газар (NOAA)-ын Үндэсний геофизикийн өгөгдлийн төв (NGDC)-өөс гаргасан Дэлхийн гадаргын рельефийн нэгдсэн (ETOPO) загварын хуурай газрын далайн түвшнээс дээших үнэмлэхүй өндөр (өндөршил)-ийн [23] болон Дэлхийн хүн амын нягтшилын өгөгдлийг [24] тус тус ашиглав.

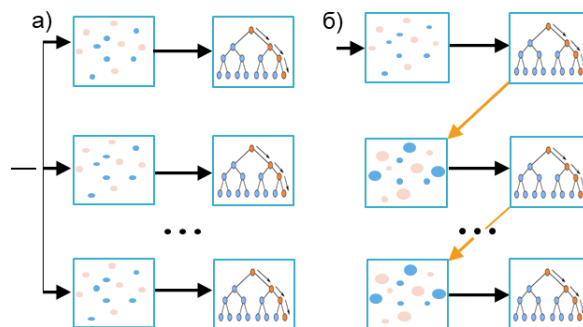
Загварт ашиглагдах 0.1° -аас бага орон зайн шийдтэй өгөгдлүүдийг хос шугаман интерполяцийн аргаар дахин тооцоолж 0.1° орон зайн шийд рүү шилжүүлсэн, мөн цаг болон өдрийн хугацааны шийд бүхий өгөгдлүүдийг 16 өдрийн арифметик дунджаар авч үл хамаарах хувьсагчдыг бэлэн болгосон. Эдгээр бэлдсэн хувьсагчид нь тоон утга, хэмжилтийн нэгжийн хувьд харилцан адилгүй байсан тул утгын хэлбэлзэл, хэмжилтийн нэгжийн зөрүүг арилгах зорилгоор хамгийн их болон бага утгаар нь нормчилсон болно.

3. АРГА ЗҮЙ

3.1. Загварын тодорхойлолт

Энэхүү судалгаанд машин сургалтын ансамблийн уутлах болон градиент бүүст сургалтын техникүүдээр дээр суурилан сурдаг регрессийн алгоритмууд (RF, XGB, LGBM) болон Stk загварыг ашигласан. Ансамбль гэдэг нь хэд хэдэн таамаглах загваруудыг нэгтгэн алдааг багасгадаг нэгдмэл сургалтын техник юм. Ансамбль алгоритмуудын сургалт нь шийдвэрийн модны шийдлийн түвшингүүдийг уутлах (bagging) болон градиент бүүст (gradient boosting) гэсэн 2 техникийг хэрэгжүүлдэг [25], [26]. Ансамблийн уутлах техник нь өгөгдлүүдийг санамсаргүй байдлаар дахин түүвэрлэх (bootstrap) замаар олон

түүвэр үүсгэн сурч, загварчлан эцэст нь эдгээр загваруудын гүйцэтгэлийг нэгтгэн эцсийн шийдийг дунджаар нь авдаг [27]. Харин градиент бүүст сургалтын техник нь функцийн градиент ашиглан алдааг багасгадаг. **Зураг 2**-д ансамблийн уутлах болон градиент бүүст сургалтын техникийн ялгааг харуулав [25], [27].



Зураг 2. Ансамблийн а) уутлах б) градиент бүүст арга

RF алгоритмыг Брэймэн 2001 онд анх хөгжүүлсэн [28]. Бидний ашигласан ансамблийн RF нь уутлах техник дээр суурилсан алгоритм юм. Харин XGB, LGBM алгоритмууд нь ансамблийн градиент бүүст техник дээр суурилан сурдаг. Өөрөөр хэлбэл функцийн градиент ашиглан өгөгдлийг шийдвэрийн модны загварууд дээр тулгуурлан алдааг дэс дараалан багасгадаг.

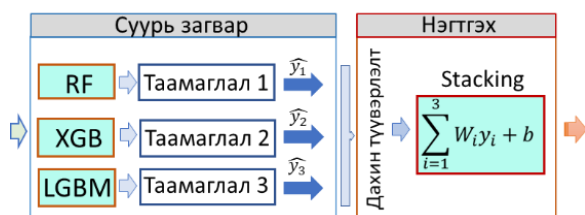
XGB нь бусад алгоритмуудтай харьцуулахад түүврийн их, бага хэмжээнээс хамаарахгүйгээр сайн гүйцэтгэлтэйгээр загварчилдаг машин сургалтын алгоритм юм [29], [30].

LGBM нь функцийн градиент ашиглан сургалтын өгөгдлийг хурдан, өндөр бүтээмжтэй боловсруулдаг төдийгүй бусад алгоритмаас сургалтын хурд өндөр, санах ойн зарцуулалт бага байдаг онцлогтой [31].

Ансамблийн уутлах болон градиент бүүст техник дээр суурилсан алгоритмууд нь загварын тогтворгүй байдал, хэвийлтийг багасгахын тулд нэг төрлийн загваруудыг нэгтгэдэг бол ансамблийн Stacking загвар нь нэгэн төрлийн биш загваруудыг нэгтгэн загварчилдаг [32]. Stk загвар нь суурь загвар, нэгтгэн дахин загварчлах гэсэн 2 үе шаттай. Бид RF, XGB, LGBM алгоритмуудаар загварчилсан үр дүнг нэгтгэн дахин түүвэрлэн авч олон хүчин зүйлийн шугаман регрессийн дээр суурилан Stk загварыг ажиллуулсан. Загварчлах явцыг Зураг 3-д дүрслэн харууллаа.

Аливаа загварт үл хамаарах хувьсагчдыг оновчтой сонгох, загварыг тайлбарлах нь чухал юм. Гэвч машин сургалтын загвар нь хар хайрцаг

мэт далд процесс тул загварыг тайлбарлах боломжгүй байдаг. Харин үнэлэгдсэн загварын үр дүнд тулгуурлан хувьсагчдын загварт оруулсан хувь нэмэр, ач холбогдлыг дүрслэн харуулдаг хиймэл оюун ухааны зарчим дээр суурилсан Таамаглалыг тайлбарлах нэгдсэн тогтолцоо Shapley Additive Explanations (SHAP) нь энэ асуудлыг шийдсэн [33]. Бид судалгаандаа хамгийн өндөр гүйцэтгэлтэй загварт тулгуурлан SHAP утгыг тооцож, үл хамаарах хувьсагчдын загварт оруулсан хувь нэмрийг үнэлж, цаашлаад CO₂-д нөлөөлөх хүчин зүйлийн уялдаа холбоо, физик процессыг тайлбарлах оролдлогыг хийлээ.



Зураг 3. Stk загварын бүтэц

3.2. Загварчлах процесс болон үнэлэх

Энэхүү судалгаанд GOSAT L2 бүтээгдэхүүний 2010-2022 оны өгөгдлийг 16 өдрийн, 0.1° торлолд дунджилсан 38891 мөр мэдээлэл бүхий XCO₂-г хамаарах хувьсагчаар, тэдгээрт харгалзах 12 үл хамаарах хувьсагчийг авч машин сургалтын 4 өөр алгоритмаар (RF, XGB, LGBM, Stk)-ыг он тус бүр дээр сургаж, загварын гүйцэтгэлийг үнэлэн, харьцуулсан. XCO₂ загварын ерөнхий бүтцийг дараах байдлаар илэрхийлж болно.

$$XCO_{2j} = f(sp_j, blh_j, rad_j, \dots) \quad (1)$$

Үүнд: XCO_{2j} нь судалгааны тухайн хугацааны торлолын нүдэнд харгалзах таамагласан утга, f функц нь торлолын тухайн байршилд XCO₂ загварчлах алгоритм (RF, XGB, LGBM, Stk), $sp_j, blh_j, rad_j, \dots$ нь загварын оролтын үл хамаарах хувьсагчид.

Загварын гүйцэтгэлийг дараах 3 хэлбэрээр шалгасан. Үүнд:

1. Загварын гүйцэтгэлийг үнэлэхдээ солбин бататгах (cv) аргаар шалгасан. Энд загварчлалын итерац бүр дээр сургалтын өгөгдлийг 10 тэнцүү хэсэгт хувааж, 9 хэсгээр сургаж, үлдсэн 1 хэсгээр загварын гүйцэтгэлийг шалган үнэлдэг. Загварын сургалтын гүйцэтгэлийг он тус бүр дээр детерминацийн коэффициент (R^2), дундаж абсолют алдаа (MAE)

болон дундаж квадрат алдаа (RMSE)-г тооцон үнэлсэн. Харин солбин бататгах аргад суурилсан гүйцэтгэлийг $R^2(cv)$ -оор үнэллээ.

2. Өндөр гүйцэтгэлтэй сурсан загварыг ашиглан таамагласан XCO₂ болон GOSAT XCO₂ хоорондох хамаарлыг шугаман регрессийн загвараар тооцон, үнэлсэн.
3. Түүнчлэн таамагласан XCO₂ болон OCO-2 хиймэл дагуулын XCO₂ өгөгдлийн хоорондох хамаарлыг шугаман регрессийн загвараар тооцон, үнэлсэн. OCO-2 XCO₂ нь 2015-2020 оны сар бүрийн дундаж XCO₂ тархалтын зураг бөгөөд 0.5°x0.625° орон зайн шийдтэй. Иймээс таамагласан XCO₂ утгыг хос шугаман интерполяцийн аргаар 0.5°x0.625° орон зайн шийд рүү хөрвүүлж, судалгааны талбайн нийт пикселийн дунджаар үнэлгээг гүйцэтгэлээ.

4. ҮР ДҮН, ХЭЛЭЛЦҮҮЛЭГ

4.1. Загварын гүйцэтгэл

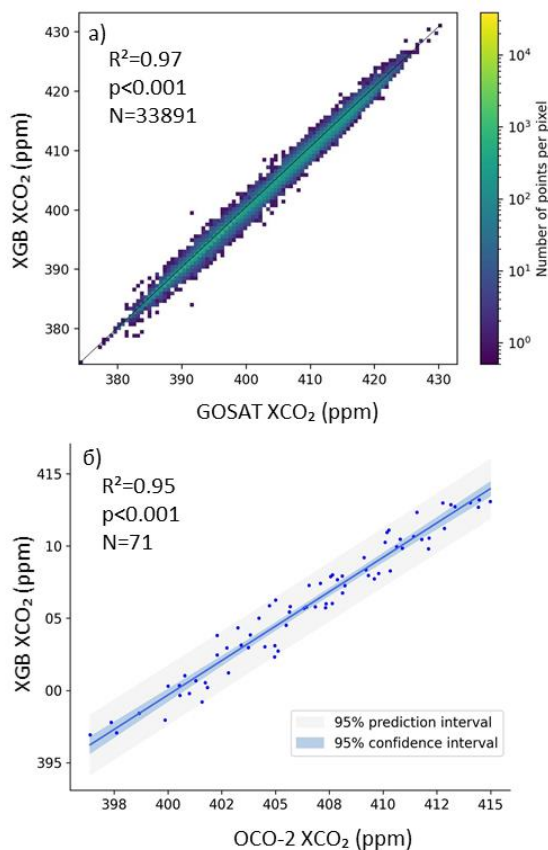
Бид XCO₂-г машин сургалтын RF, XGB, LGBM регрессийн загварууд болон Stk загваруудаар он тус бүрээр загварчлан, гүйцэтгэлийг үнэлсэн (гүйцэтгэлийг энэхүү өгүүлэлд харуулаагүй). Загваруудын R^2 нь 0.76-0.98 хооронд, MAE алдаа нь 0.19-1.29 ppm хооронд, RMSE нь 0.48-1.68 ppm алдаатайгаар сурсан бол солбин бататгах арга дээр суурилсан шалгах өгөгдлийн R^2 (cv) нь 0.44-0.96 хооронд байсан. **Зураг 4**-д судалгааны хугацааны он тус бүр дээр тооцсон загваруудын алдааны дунджийг харуулав.

	MAE (ppm)	RMSE (ppm)	R^2	R^2 (cv)
RF	0.71	0.94	0.92	0.65
XGB	0.36	0.62	0.97	0.73
LGBM	0.67	0.96	0.91	0.62
Stk	1.1	1.4	0.83	0.61

Зураг 4. Загваруудын гүйцэтгэлийн дундаж

Загвар тус бүрээр гүйцэтгэлийг авч үзвэл MAE нь Stk>RF>LGBM>XGB буюу XGB загвараар таамагласан XCO₂ утга, бодит утгын хоорондох абсолют зөрүү хамгийн бага байна. RMSE хувьд Stk>LGBM>RF>XGB буюу мөн адил XGB загвар хамгийн өндөр гүйцэтгэлийг үзүүлсэн. Сургалтын өгөгдлийн R^2 нь Stk<LGBM<RF<XGB буюу XGB загвар нь XCO₂ утгыг хамгийн өндөр тайлбарлах чадвартайгаар сурсан. Харин $R^2(cv)$ хувьд Stk<LGBM<RF<XGB байна. Эндээс XGB загвар нь хамгийн өндөр гүйцэтгэлтэй буюу загварт ороогүй XCO₂ өгөгдлийн 73%-ийг дунджаар

тайлбарлах чадвартай байлаа. XGB загварын гүйцэтгэлийг он тус бүрээр тооцсон үр дүнгээс авч үзвэл сургалтын өгөгдлийн хувьд $MAE < 0.61$ ppm, $RMSE < 0.83$ ppm, $R^2 > 0.95$, харин $R^2(cv)$ нь 0.55-аас 0.96 хооронд байсан. Иймд XGB загваруудыг ашиглан 2010 оноос 2022 он хүртэлх хугацааны 16 өдрийн, 0.1° орон зайн шийдтэй XCO_2 зургийг гаргасан. Загвараар таамагласан XGB XCO_2 болон пиксельд харгалзах GOSAT XCO_2 утгуудын хамаарлыг шугаман регрессийн загвараар үнэлснийг Зураг 5а-д харуулав. Энд хамаарал нь $R^2 = 0.98$, $p < 0.001$ буюу хангалттай сайн үр дүнг үзүүлжээ. Түүнчлэн OCO-2 XCO_2 өгөгдлийн 2015-2020 оны сар бүрийн нийт пикселийн дундаж утга болон таамагласан XGB XCO_2 хоорондох хамаарал ($R^2 = 0.95$, $p < 0.001$) мөн адил маш сайн гарсан (Зураг 5б).



Зураг 5. XGB XCO_2 болон а) GOSAT XCO_2 , б) OCO-2 XCO_2 хамаарал

4.2. XCO_2 агууламжид нөлөөлөх хүчин зүйл

XGB XCO_2 загварт тулгуурлан 2010-2022 хугацааны он тус бүр дээр SHAP утгыг тооцсон (энэхүү өгүүлэлд он тус бүр тооцсон үр дүнг харуулаагүй). SHAP утгаас харахад үл хамаарах

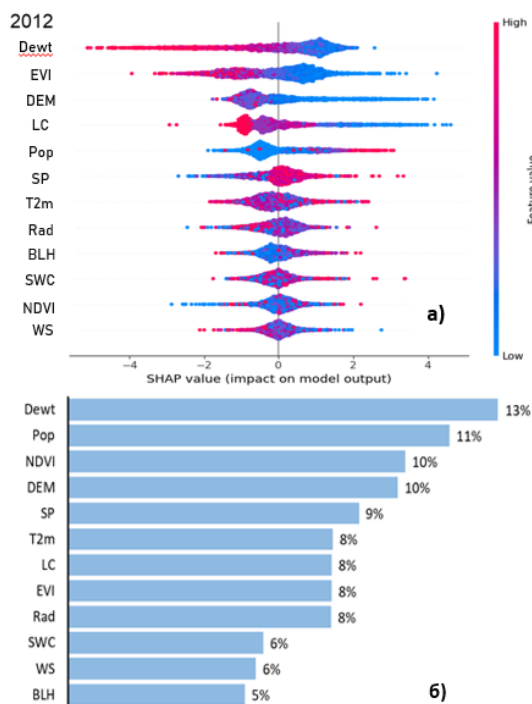
хувьсагчдын загварт үзүүлэх нөлөө, ач холбогдол харилцан адилгүй гарсан.

Жишээ болгон 2012 он дээрх SHAP утгын тархалтыг **Зураг 6а**-д үзүүлэв. Энд босоо тэнхлэг нь SHAP утгыг, хэвтээ тэнхлэг нь загварт үзүүлэх эерэг, сөрөг нөлөөг илэрхийлнэ. 2012 оны хувьд загварт шүүдэр цэгийн температур (Dewt), ургамлын сайжруулсан индекс (EVI), гадаргын өндөршил (DEM)-ийн их утга (улаан өнгө) XCO_2 таамаглалын утгыг бууруулж, сөрөг нөлөө үзүүлжээ. Энд усны уурын агууламж их байхад ургамлын фотосинтезийн процесс эрчимжиж CO_2 шингээх чадавх нэмэгдсэнээр загварын XCO_2 таамаглалын утгыг бууруулах нөлөө үзүүлсэн. Харин нам өндөрт (DEM) байрлах нутаг оронд CO_2 ялгарлын эх үүсвэрээс шалтгаалсан хуримтлал үүссэн байх боломжтой. Агаарын даралт (SP)-ын хувьд их утга нь загварын XCO_2 таамаглалын утгыг нэмэгдүүлж, бага утга (хөх өнгө) нь бууруулсан. Өөрөөр хэлбэл гадарга орчмын агаарын даралт нь XCO_2 -тай эерэг хамааралтай байна.

Үүнийг их даралтын орон нь тухайн газар нутагт агаарын масс тогтвортой байх нөхцлийг бүрдүүлж, улмаар CO_2 агууламж ихтэй агаарын масс хуримтлагд-санаар XCO_2 таамаглах утгыг нэмэгдүүлсэн гэж тайлбарлаж болно. Хамгийн бага нөлөөлсөн WS хувьд их, бага утга босоо тэнхлэгийн 2 талд бараг адил тархаж, их утга нь XCO_2 таамаглалын утгыг улам нэмэгдүүлсэн бол, бага утга нь бууруулжээ. Эндээс 2012 онд WS үл хамаарах хувьсагч нь XCO_2 -тай загварын ихэнхи тохиолдолд урвуу хамааралтай байсан гэж дүгнэж болно. Бусад үл хамаарах хувьсагчдын хувьд SHAP утгын тархалтын ихэнх нь босоо тэнхлэгт ойрхон тархаж загварт сул, эерэг, сөрөг хам нөлөө үзүүлжээ.

2010-2022 хугацааны он тус бүр дээр тооцсон SHAP утгын дундаж тархалтыг хувиар илэрхийлэн **Зураг 6б**-д үзүүлэв. Энд $Dewt > Pop > NDVI > DEM > SP > T2m > LC > EVI > Rad > SWC > WS > BLH$ буюу судалгааны хугацаанд шүүдэр цэгийн температур хамгийн их (13%), хязгаарын үе давхаргын өндөр хамгийн бага (5%) нөлөө үзүүлжээ. Dewt нь агаарт агуулагдах усны уурын агууламжийг илэрхийлэх хэмжигдэхүүн юм. Агаар дахь хийн хольцуудын нэг агууламжийн хэмжээ өөрчлөгдөхөд бусад агууламжийн хэмжээ шахагдах (багасах) эсвэл хуримтлагдана. Иймд XGB XCO_2 таамаглах загварт Dewt нь хамгийн их буюу 13%-ийн нөлөө үзүүлсэн гэж үзэж байна. Хүн амын нягтшил (Pop) нь CO_2 ялгарлын эх үүсвэртэй холбоотой хувьсагч бөгөөд таамаглах загварт мөн адил хүчтэй нөлөөлсөн. Ургамлын фотосинтезийн процесстэй

холбоотойгоор ургамлын индекс (NDVI, EVI)-үүд нь биосфер-агаар мандлын хоорондох нүүрстөрөгчийн эргэлтийг тодорхойлогч гол индикатор болдог. Иймээс XGB XCO₂ загварт чухал хувьсагч болж байна.



Зураг 6. Үл хамаарах хувьсагчдын загварт үзүүлсэн нөлөөллийн а) утгын тархалт б) харьцангуй ач холбогдол

Харин BLH нь газрын гадаргын атираашил, салхины хурд болон дулааны горимын тогтворгүй байдлаас хамаардаг төдийгүй агаарын бохирдуулагч хольцын агууламжид хүчтэй нөлөөлдөг. Өөрөөр хэлбэл BLH нь агаарын турбулент солилцоог илэрхийлэгч нэг хүчин зүйл бөгөөд агаарт хийн хольц хуримтлагдах, эсвэл сарнихад хам нөлөө үзүүлж байдаг.

XGB XCO₂ загварт үзүүлсэн нийт нөлөөллийг цаг уурын хэмжигдэхүүн, ургамлын индекс, бусад (DEM, LC) болон хүний нөлөөлөл гэсэн 4 категориор авч үзэхэд харгалзан 53%, 18%, 18%, 11% байна.

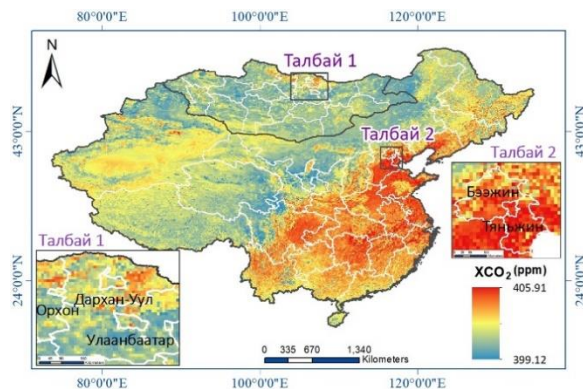
4.3. XGB XCO₂ агууламжийн орон зайн тархалт

Өндөр гүйцэтгэлтэй ажилласан машин сургалтын XGB загварыг ашиглан 2010-2022 оны 16 өдрийн, 0.1°x0.1° орон зайн шийдтэй нийт 296 (энд 2015 оны 001, 017, 033-р өдөр тооцоонд ороогүй) XGB XCO₂ орон зайн тархалтын зургийг

гаргалаа. Орон зайн тархалтаас 2010 оны 193-р өдрийн (VII.12) агууламж хамгийн бага тархалттай буюу 376.4-397.8 ppm хооронд, судалгааны талбайн нийт 113706 пикселийн дунджаар 386.8±1.6 ppm байсан бол хамгийн их XGB XCO₂ нь 2022 оны 049-р өдрийн (II.18) тархалтын зурагт 408.8-427.8 ppm хооронд буюу дундаж утга нь 417.1±1.6 ppm байлаа.

XGB XCO₂ орон зайн тархалтын 2010-2022 оны дунджийг **Зураг 7**-д харуулав. XGB XCO₂ тархалтын зургаас үзэхэд их утга БНХАУ-ын зүүн болон зүүн өмнөд бүсийн хүн ам, эдийн засаг төвлөрсөн бүс нутгуудаар ажиглагдсан бол бусад хэсгээр ялангуяа Төвөд болон Монголын тэгш өндөрлөгөөр бага утга ажиглагдсан. Бидний загварчилсан XGB XCO₂ орон зайн тархалтын энэ хэв маяг олон судлаачдын үр дүнтэй адил байна.

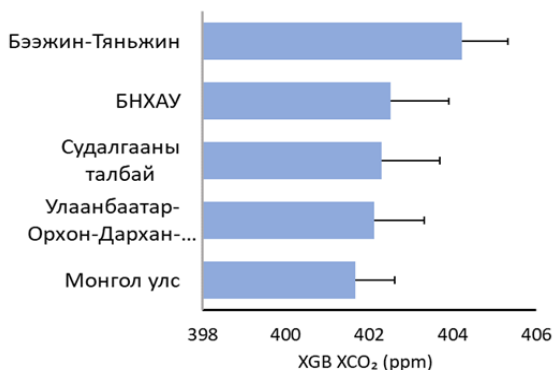
Тухайлбал, Zhao ба бусад (2019) судалгаандаа XCO₂ агууламжийн их утга Дэлхийн хойд хагаст, ялангуяа БНХАУ-ын зүүн эргийн нутгаар тархаж байгааг дурдсан. Ялангуяа Төвөдийн тэгш өндөрлөг, Алтайн нуруу, Монголын тэгш өндөрлөг зэрэг өндөрлөг газруудад бүтэн жилийн турш CO₂ бага утга ажиглагдаж байгааг тогтоосон [34]. Энэ дүгнэлт нь бидний загварт гадаргын өндөршил (Dem) нь чухал нөлөөлөгч гэдгийг баталж байна (**Зураг 6**).



Зураг 7. 2010-2022 оны XGB XCO₂ дундаж утгын тархалт

2010-2022 оны XGB XCO₂ дундаж утгын тархалтыг нийт пикселийн дунджаар авч үзвэл Монгол улс 401.7±0.9 ppm, Улаанбаатар хот, Орхон, Дархан-Уул аймгийг хамарсан талбайд 402.13±1.2 ppm буюу хамгийн бага (**Зураг 8**) байлаа. Харин судалгааны нийт талбайд 402.3±1.4 ppm байна. БНХАУ-ыг хамарсан талбайд 402.51±1.4 ppm буюу Монгол улсыг хамарсан талбайн дунджаас 0.84 ppm-ээр, судалгааны нийт талбайн дунджаас 0.62 ppm-ээр тус тус их байна. БНХАУ-ын Бээжин, Тяньжин хотыг хамарсан

талбайн дундаж 404.21 ± 1.1 ppm байгаа нь БНХАУ-ыг хамарсан талбайгаас 1.70 ppm-ээр, судалгааны нийт талбайгаас 1.92 ppm-ээр тус тус их байна. Эндээс үзэхэд Монгол улс, БНХАУ-ын хүн ам төвлөрсөн, эдийн засгийн хөгжсөн газар нутагт CO_2 агууламжийн их утга тархдаг [35], [36] гэсэн судалгаануудтай адил байна.

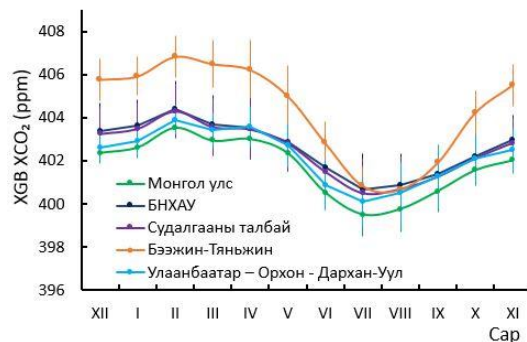


Зураг 8. 2010-2022 оны XGB CO_2 дундаж утгын харьцуулалт, талбайгаар

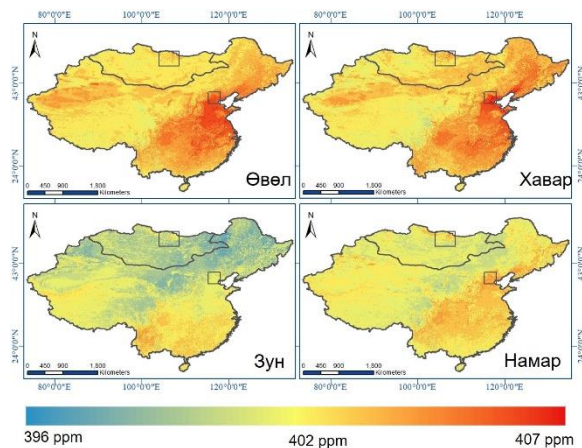
4.4. XGB CO_2 агууламжийн улирлын орон зайн тархалт

2010-2022 оны сарын дундаж XGB CO_2 орон зайн тархалтын жилийн доторх хуваарилалтыг талбайд харгалзах пикселийн дундаж утгаар нь Зураг 9-д харууллаа. Зургаас үзэхэд XGB CO_2 агууламж нь өвлийн саруудад их утгаа хадгалж хаврын улирлаас буурч зун хамгийн бага утгадаа хүрээд намрын улиралд буцаад өсдөг зүй тогтол ажиглагдаж байна. Энэ нь өвлийн улиралд азийн эсрэг циклон тогтворжиж [37] хүйтний улиралд эрчим хүчний хэрэглээ, галлагаанаас шалтгаалсан CO_2 агууламжийн их утга удаан хадгалагдах нөхцөл бүрддэг. Хаврын улиралд аажмаар CO_2 буурдаг байна. Зургаас үзэхэд, IV сард бага зэргийн XGB CO_2 өсөлт ажиглагдсан. Энэ нь хавар хөрсний гадарга орчмын температур өсч дэвсгэр гадаргын чийг дулаан солилцоо эрчимжиж, хөрсөнд агуулагдах нүүрстөрөгч агаарт гарах нөхцөл бүрддэгтэй холбоотойгоос гадна судлаачдын үзэж буйгаар тариалалт эхлэж, хөрс боловсруулалтаас үүдэн агаар мандалд гарах CO_2 өсдөгтэй холбоотой [38]. V сараас буурч VII, VIII сард хамгийн бага агууламж ажиглагдсан нь CO_2 -г шингээх ургамлын фотосинтезийн процесс давамгайлсантай холбоотой. Харин намрын улиралд фотосинтезийн процесс аажмаар буурч ургамлын амьсгал давамгайлж эхэлдэг. Талбайгаар нь харьцуулбал Бээжин, Тяньжин хотуудыг хамарсан талбайд бусад талбайтай

харьцуулахад бараг бүх саруудад их CO_2 агууламжтай байна.



Зураг 9. XGB CO_2 жилийн доторх хуваарилалт, 2010-2022



Зураг 10. 2010-2022 оны XGB CO_2 дундаж улирлын тархалт

Зураг 10-д 2010-2022 оны XGB CO_2 дундаж улирлын орон зайн тархалтыг үзүүлээ. Зургаас үзэхэд XGB CO_2 агууламжийн тархалтын их утга өвөл $389.9\text{--}407.5$ ppm хооронд, нийт пикселийн дунджаар 403.4 ± 1.2 ppm, хавар $398.9\text{--}407.9$ ppm, дунджаар 403.3 ± 1.3 ppm байсан бол бага агууламж зун $396.5\text{--}405.7$ ppm хооронд, дунджаар 400.9 ± 1.5 ppm, намар $398.5\text{--}406.1$ ppm, дунджаар 402.1 ± 1.0 ppm байна. Орон зайн тархалтын хувьд мөн адил бүх улиралд БНХАУ-ын зүүн болон зүүн өмнөд бүсийн хүн ам, эдийн засаг төвлөрсөн бүс нутгуудаар ажиглагдсан бол бусад хэсгээр ялангуяа Төвөд болон Монголын тэгш өндөрлөгөөр бага утга ажиглагдсан. Орон зайн тархалтын энэ зүй тогтол бусад судлаачдын судалгаатай ижил байна. Тухайлбал, Zhao ба бусад (2019) судалгаандаа зуны улиралд хүчтэй зуны муссоны нөлөөнд ялангуяа БНХАУ-ын зүүн өмнөд бүсийн нутаг дэвсгэрт CO_2 агууламж ихтэй агаарын масс хуримтлагдах нөхцөл бүрддэг гэжээ. Харин өндөрлөг газар, ялангуяа Төвөд, Монголын

тэгш өндөрлөгт муссоны чиглэл өөрчлөгдөхөөс гадна салхины хэвтээ чиглэл дэх урсгал хүчтэй болж салхи нь сарниулагчийн үүрэг гүйцэтгэн, CO₂ бага агууламж ажиглагдах нөхцөлийг бий болгодог. Өвлийн улиралд баруун, баруун хойноос чиглэлтэй салхи зонхилж БНХАУ-ын зүүн, зүүн хойд хэсэгт CO₂ агууламж ихтэй агаарын масс хуримтлагдах нөхцөл бүрдэхээс гадна, хүний үйл ажиллагаатай холбоотой эрчим хүчний хэрэглээ агууламжийг улам нэмэгдүүлдэг байна [34].

5. ДҮГНЭЛТ

Энэхүү судалгаандаа Монгол, БНХАУ-ын нутаг дэвсгэрийг хамарсан агаар мандал дахь XCO₂-г машин сургалтын ансамблийн RF, XGB, LGBM регрессийн алгоритмууд болон Stk загвараар загварчилж харьцууллаа. Загваруудаас функцийн градиент ашиглан алдааг багасган сурдаг техник дээр суурилсан алгоритмууд (XGB, LGBM) өндөр гүйцэтгэлтэй сурсан. Судалгааны энэ үр дүн нь өмнөх судлаачдын үр дүнтэй адил байна [9], [10]. Хамгийн өндөр гүйцэтгэлтэй загвар нь XGB ($R^2=0.97$, $R^2(cv)=0.73$) байлаа.

Түүнчлэн өндөр гүйцэтгэлтэй загварчилсан XGB XCO₂ үр дүнд тулгуурлан SHAP утгыг тооцож XCO₂-д нөлөөлөх хүчин зүйлийн уялдаа холбоо, физик процессыг тайлбарласан. Загварт үзүүлсэн нийт нөлөөллийг цаг уурын хэмжигдэхүүн, ургамлын индекс, бусад (DEM, LC) болон хүний нөлөөлөл гэсэн 4 категориор авч үзэхэд харгалзан 53%, 18%, 18%, 11% -иар загварт нөлөөлсөн.

Судалгааны үр дүн нь уур амьсгалын хэлбэлзлийг судлах, CO₂ ялгарлын эх үүсвэрийг илрүүлэх, биосферийн нүүрстөрөгчийн эргэлтийг загварчлахад суурь өгөгдөл болох боломжтой. Цаашдаа илүү өндөр нарийвчлалтай зайнаас тандан судлалын өгөгдлөөр судалгааны ажлыг сайжруулах шаардлагатай.

ТАЛАРХАЛ

Судалгааны ажлыг хийж гүйцэтгэхэд тусалж дэмжсэн Газарзүй, Геоэкологийн хүрээлэнгийн хамт олон, найз нөхд болон үнэтэй зөвлөгөө өгсөн доктор А.Саруулзаяад талархал илэрхийлье.

АШИГЛАСАН МАТЕРИАЛ

- [1] P. Friedlingstein *et al.*, “Global Carbon Budget,” *Earth Syst. Sci. Data*, vol. 14, pp. 4811–4900, 2022, <https://doi.org/10.5194/essd-14-4811-2022>.

- [2] IPCC, “Global Carbon and Other Biogeochemical Cycles and Feedbacks,” in *Climate Change 2021 – The Physical Science Basis*, I. P. on C. C. (IPCC), Ed., Cambridge: Cambridge University Press, 2023, pp. 673–816. Available: <https://doi.org/10.1017/9781009157896.007>.
- [3] Д. Дагвадорж, *Агаар мандлын физик-химийн судалгаа*. Улаанбаатар, 2023.
- [4] M. Zhang and G. Liu, “Mapping contiguous XCO₂ by machine learning and analyzing the spatio-temporal variation in China from 2003 to 2019,” *Sci. Total Environ.*, vol. 858, p. 159588, 2023. Available: <https://doi.org/10.1016/j.scitotenv.2022.159588>.
- [5] H. Yang, T. Li, J. Wu, and L. Zhang, “Inter-comparison and evaluation of global satellite XCO₂ products,” *Geo-spatial Inf. Sci.*, vol. 28, no. 1, pp. 131–144, Jan. 2025. Available: <https://doi.org/10.1080/10095020.2023.2252017>.
- [6] R. Doughty *et al.*, “Global GOSAT, OCO-2, and OCO-3 solar-induced chlorophyll fluorescence datasets,” *Earth Syst. Sci. Data*, vol. 14, no. 4, pp. 1513–1529, 2022. Available: <https://doi.org/10.5194/essd-14-1513-2022>.
- [7] F. Kataoka, D. Crisp, T. E. Taylor, C. W. O'Dell, A. Kuze, K. Shiomi, H. Suto, C. Bruegge, F. M. Schwandner, R. Rosenberg, et al., “The Cross-Calibration of Spectral Radiances and Cross-Validation of CO₂ Estimates from GOSAT and OCO-2,” *Remote Sens.*, vol. 9, p. 1158, 2017. doi: 10.3390/rs9111158. C. Rakauckas, A. Edelman, and R. Ferrari, “A Modular System for Increasing Predictiveness for Extreme Climate Predictions,” United States, 2021, Available: <https://doi.org/10.2172/1769647>.
- [8] S. He, Y. Yuan, Z. Wang, L. Luo, Z. Zhang, H. Dong, and C. Zhang, “Machine Learning Model-Based Estimation of XCO₂ with High Spatiotemporal Resolution in China,” **Atmosphere**, vol. 14, p. 436, 2023, <https://doi.org/10.3390/atmos14030436>.
- [9] C. He *et al.*, “Deriving Full-Coverage and Fine-Scale XCO₂ Across China Based on OCO-2 Satellite Retrievals and CarbonTracker Output,” *Geophys. Res. Lett.*, vol. 49, no. 12, 2022. Available: <https://doi.org/10.1029/2022GL098435>.

- [10] J. Li, K. Jia, X. Wei, M. Xia, Z. Chen, Y. Yao, X. Zhang, H. Jiang, B. Yuan, G. Tao, and L. Zhao, "High-spatiotemporal resolution mapping of spatiotemporally continuous atmospheric CO₂ concentrations over the global continent," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 108, p. 102743, Apr. 2022. Available: <https://doi.org/10.1016/j.jag.2022.102743>.
- [11] А. Саруулзаяа ба Д. Сайнбаяр, "Нүүрсхүчлийн хий (CO₂) болон метан (CH₄) хийн агууламжийн орон зайн тархалт," *Монгол Орны Газарзүй-Геоэкологи Сэтгүүл - МОГТС*, vol. 42, pp. 252–259, 2022.
- [12] Д. Сайнбаяр, С. Эрдэнэсүх, ба А. Саруулзаяа, "Монгол орны нүүрсхүчлийн хийн хэлбэлзлийн судалгаа," *Geogr. Issues*, vol. 22, no. 1, pp. 45–56, Mar. 2022. [Online]. Available: <https://journal.num.edu.mn/GP/article/view/536>.
- [13] S. Adiya, S. Dalantai, T. Wu, X. Wu, J. Yamkhin, Y. Bao, E. Sumiya, G. Yadamsuren, D. Avirmed, and B. Dorjgotov, "Spatial and temporal change patterns of near-surface CO₂ and CH₄ concentrations in different permafrost regions on the Mongolian Plateau from 2010 to 2017," *Sci. Total Environ.*, vol. 800, p. 149433, Dec. 2021. Available: <https://doi.org/10.1016/j.scitotenv.2021.149433>.
- [14] S. Yang, X. Wang, Z. Ge, G. Dong, M. Ma, and X. Han, "Global Per Capita CO₂ Emission Trends," 2023. Available: <https://doi.org/10.3390/atmos14121797>.
- [15] F. P. Lopez, G. Zhou, G. Jing, K. Zhang, and Y. Tan, "XCO₂ and XCH₄ Reconstruction Using GOSAT Satellite Data Based on EOF-Algorithm," 2022. Available: <https://doi.org/10.3390/rs14112622>.
- [16] T. Yokota *et al.*, "Global Concentrations of CO₂ and CH₄ Retrieved from GOSAT: First Preliminary Results," *SOLA*, vol. 5, pp. 160–163, 2009. Available: <https://doi.org/10.2151/sola.2009-041>.
- [17] T. A. Adjuik and S. C. Davis, "Machine Learning Approach to Simulate Soil CO₂ Fluxes under Cropping Systems," 2022. Available: <https://doi.org/10.3390/agronomy12010197>.
- [18] X. Chen, Q. He, T. Ye, Y. Liang, and Y. Li, "Decoding spatiotemporal dynamics in atmospheric CO₂ in Chinese cities: Insights from satellite remote sensing and geographically and temporally weighted regression analysis," *Sci. Total Environ.*, vol. 908, p. 167917, 2024. Available: <https://doi.org/10.1016/j.scitotenv.2023.167917>.
- [19] J. Muñoz-Sabater, "ERA5-Land hourly data from 1950 to present," 2019, *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. Available: <https://doi.org/10.24381/cds.e2161bac>.
- [20] H. Boogaard, J. Schubert, A. De Wit, J. Lazebnik, R. Hutjes, and G. Van der Grijn, "Agrometeorological indicators from 1979 to present derived from reanalysis," *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, 2020, <https://doi.org/10.24381/cds.6c68c9bb>.
- [21] H. Hersbach *et al.*, "ERA5 hourly data on single levels from 1940 to present," *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, 2023, <https://doi.org/10.24381/cds.adbb2d47>.
- [22] NOAA National Geophysical Data Center, "2-minute Gridded Global Relief Data (ETOPO2) v2. NOAA National Centers for Environmental Information".
- [23] A. J. Tatem, "WorldPop, open data for spatial demography," *Sci. Data*, vol. 4, no. 1, p. 170004, 2017. Available: <https://doi.org/10.1038/sdata.2017.4>.
- [24] T. Maganathan, S. Senthilkumar, and V. Balakrishnan, "Machine Learning and Data Analytics for Environmental Science: A Review, Prospects and Challenges," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 955, no. 1, 2020. Available: <https://doi.org/10.1088/1757-899X/955/1/012107>.
- [25] P. Xu, "Review on Studies of Machine Learning Algorithms," *J. Phys. Conf. Ser.*, vol. 1187, no. 5, p. 52103, 2019. Available: <https://doi.org/10.1088/1742-6596/1187/5/052103>.
- [26] B. Naveen and S. Manu, "Detecting DDoS Attacks Using Machine Learning Techniques and Contemporary Intrusion Detection Dataset," *Autom. Control Comput. Sci.*, vol. 53, no. 5, pp. 419–428, 2019.

- <https://doi.org/Available:10.3103/S0146411619050043>.
- [27] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, p. 5, 2001.
- [28] C. Zheng, A. Abd-Elrahman, V. Whitaker, and C. Dalid, "Prediction of Strawberry Dry Biomass from UAV Multispectral Imagery Using Multiple Machine Learning Methods," 2022. Available: <https://doi.org/10.3390/rs14184511>.
- [29] N. M. Shahani, X. Zheng, C. Liu, F. U. Hassan, and P. Li, "Developing an XGBoost Regression Model for Predicting Young's Modulus of Intact Sedimentary Rocks for the Stability of Surface and Subsurface Structures," *Front. Earth Sci.*, vol. 9, no. October, pp. 1–13, 2021. Available: <https://doi.org/10.3389/feart.2021.761990>.
- [30] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," Microsoft Research and Peking University, 2017. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.
- [31] M. Kuhn and K. Johnson, *Applied Predictive Modeling*, vol. 26. 2013.
- [32] S. M. Lundberg and S.-I. Lee, *A Unified Approach to Interpreting Model Predictions*. 2017.
- [33] J. Zhao, G. Li, W. Cui, Q. Cao, and H. Zhang, "Important evidence of constant low CO₂ windows and impacts on the non-closure of the greenhouse effect," *Sci. Rep.*, vol. 9, no. 1, pp. 1–13, 2019. Available: <https://doi.org/10.1038/s41598-019-41562-x>.
- [34] L. Wei, H. Liu, and L. Wu, "Spatial distribution of floating population in Beijing, Tianjin and Hebei Region and its correlations with synergistic development," *Math. Biosci. Eng.*, vol. 20, no. 3, pp. 5949–5965, 2023. Available: <https://doi.org/10.3934/mbe.2023257>.
- [35] ҮТХН, "Парисын хэлэлцээрийг хэрэгжүүлэх үндэсний хэмжээнд тодорхойлсон хувь нэмрийн зорилт," 2020, *БОАЖЯ, Улаанбаатар*. [Online]. Available: <https://www.ccrcc.mn/report>.
- [36] Д. Сандэлгэр, Г. Болорчимэг, ба С. Эрдэнэсүх, "Мөрөн хот орчмын орчих мандлын дулааны судалгаа," *Монгол Орны Газарзүй-Геоэкологи Сэтгүүл - МОГТС*, vol. Тусгай дуг, pp. 97–103, 2017.
- [37] L. Shen, C. Zhao, Z. Ma, Z. Li, J. Li, and K. Wang, "Observed decrease of summer sea-land breeze in Shanghai from 1994 to 2014 and its association with urbanization," *Atmos. Res.*, vol. 227, pp. 198–209, 2019. Available: <https://doi.org/10.1016/j.atmosres.2019.05.007>.