LATE REQUEST FOR A EMI R&D PROJECT 2026–2028

Country/Organisation:	United Kingdom
Principal Investigator[1]:	Akash Koppa
Affiliation:	University of Reading
Address:	Department of Geography and Environmental Sciences, Russell Building, University of Reading, Reading RG6 6DR, United Kingdom
Other researchers:	N/A
Project Title:	Self-intensification of vegetation greening across the phenological cycle through land–atmosphere feedbacks

To make changes to an existing project please submit an amended version of the original form.

If this is a continuation of an existing project, please state the computer project account assigned previously.	N/A	
Starting year: (A project can have a duration of up to 3 years, agreed at the beginning of the project.)	2026	
Would you accept support for 1 year only, if necessary?	YES ☒	NO ☐

Late EMI R&D Projects are normally approved for 1 year, but exceptionally can be approved for 3 years.

Computer resources required for project year:		**2026**	**2027**	**2028**
High Performance Computing Facility	[SBU]	18M		
Graphics Processing Unit Cluster-A	[GBU]			
Graphics Processing Unit Cluster-B	[GBU]			
Accumulated data storage (total archive volume)[2]	[GB]	15,000		

[1] The Principal Investigator will act as contact person for this EMI R&D Project and, in particular, will be asked to register the project, provide annual progress reports of the project's activities, etc.
[2] These figures refer to data archived in ECFS and MARS. If e.g. you archive x GB in year one and y GB in year two and don't delete anything you need to request x + y GB for the second project year etc.
[3]The number of vGPU is referred to the equivalent number of virtualized vGPUs with 8GB memory.

EWC resources required for project year:		**2026**	**2027**	**2028**
Number of vCPUs	[#]			
Total memory	[GB]			
Storage	[GB]			
Number of vGPUs[3]	[#]			

Continue overleaf.

Principal Investigator:	Akash Koppa
Project Title:	Self-intensification of vegetation greening across the phenological cycle through land–atmosphere feedbacks

Extended abstract

A. Scientific Plan

A.1. Introduction

Satellite records show a persistent and widespread greening trend across much of the global land surface over the past four decades. Most of this trend is attributed to the CO_2 fertilisation effect, with smaller contributions from nitrogen deposition, climate change, and land-cover change (Zhu et al., 2016). Greening, represented by enhanced leaf area exerts a first-order control on the exchange of water and energy at the land surface and feeds back into the atmosphere (Forzieri et al., 2020). For example, through changes in latent heat flux and moisture recycling, relative humidity and precipitation is enhanced in both the local and downwind regions (Yu et al. 2017; Koppa et al. 2026). Shifts in vegetation phenology, a consequence of the greening trend, has resulted in seasonal changes in land surface temperature (Duiveller et al. 2018). Therefore, a positive feedback loop exists between vegetation greening and key atmospheric variables that contribute to the effect: precipitation, temperature, and radiation. While the effect of greening on each of these climate variables are relatively better understood, whether vegetation greening in one region can induce more greening in the downwind direction through these land–atmosphere feedbacks is unknown.

This project aims to mechanistically quantify the prevalence and strength of this feedback – vegetation greening enhances greening in the downwind regions through changes in precipitation, temperature, and radiation – by combining an atmospheric moisture transport model with satellite-based estimates of land oceanic fluxes (both latent and sensible heat. We have shown that such a feedback operates in drylands, where the warming and drying of air passing over expanding drylands reduces downwind precipitation and drives further aridification (Koppa et al., 2024). Here the same mechanism is quantified in the opposite direction (greening rather than drying) across the entire phenological cycle. The analysis is performed over Europe, which spans a gradient from the water-limited Mediterranean to the energy-limited north.

A.2. Hypothesis

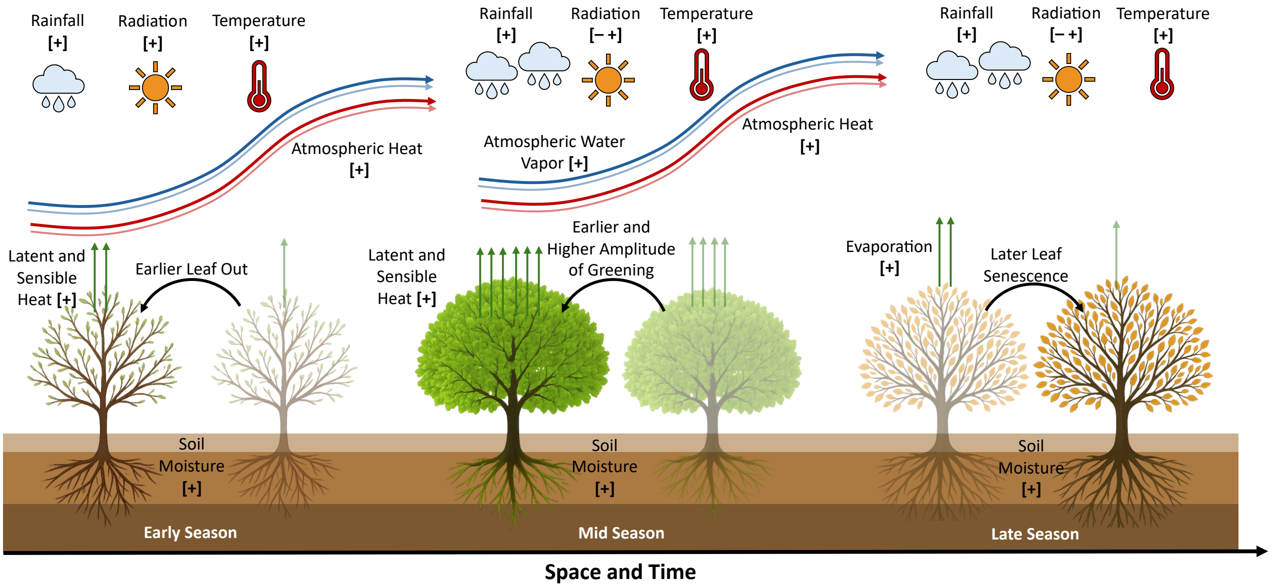

***Figure 1: Hypothesis**. Physical mechanism by which greening can self-intensify across the different stages of the vegetation phenological cycle.*

Vegetation greening is said to *self-intensify* if a positive leaf-area anomaly at the downwind (sink) region can be mechanistically attributed to enhanced leaf-area in the upwind vegetation. The mechanism is hypothesized to operate in three stages. First, a source leaf-area anomaly perturbs the surface energy balance, raising latent heat flux through greater transpiring area and sensible heat flux through reduced albedo. Second, atmospheric transport carries the resulting moisture and heat anomalies downwind to the sink. Third, these arriving anomalies modify sink phenology through two pathways: a hydrological pathway, where advected moisture enhances precipitation and a thermal pathway, where advected sensible heat and water vapour raise air temperature. Both pathways can affect the radiation at the sink by altering cloud cover and incoming shortwave radiation.

A.3. Objectives

The project has three objectives, each corresponding to one link in the hypothesised feedback loop.

Objective 1. Quantify how anomalies in vegetation greening modify surface latent and sensible heat fluxes across the phenological cycle, separating the flux response attributable to leaf area from that driven by the concurrent meteorological state.

Objective 2. Quantify the efficiency with which the moisture and heat exported by these flux anomalies are transported to downwind regions and delivered there as precipitation, water vapour, and sensible heat.

Objective 3. Quantify the downwind vegetation response to the delivered moisture and heat, combine the three links into a single source–sink feedback coefficient, and determine the sign, magnitude, and dominant pathway of the feedback across the hydroclimatic regimes of Europe.

A.4. Methodology

The hypothesis is tested by by combining Lagrangian moisture and heat tracking with satellite-observations of land and oceanic fluxes. The three physical links — the surface flux response to greening, the atmospheric transport of the resulting anomalies, and the downwind vegetation response — are estimated independently and then combined into a single coupling coefficient.

Step 1 — Lagrangian moisture and heat tracking. Air-parcel trajectories are computed with the FLEXPART model (Stohl et al., 2005) driven by ERA5. Source–receptor relationships for moisture and sensible heat are established with the HAMSTER framework (Keune et al., 2022), which diagnoses evaporation, precipitation, and sensible heat along trajectories, attributes their sources over multi-day backward trajectories, and bias-corrects the attributed fluxes against observed evaporation and precipitation. The bias-corrected output provides, for each source–sink pair and each month, the efficiency with which exported moisture reaches the sink as precipitation and as water vapour, and the efficiency with which sensible heat reaches the sink.

Step 2 — Surface flux response to greening. Monthly anomalies of latent and sensible heat flux at the source are regressed on leaf area together with the concurrent meteorological drivers. The coefficient on leaf area isolates the vegetation-driven component of each flux anomaly. The regression is estimated separately for each month to resolve the phenological cycle; the expected shift toward latent heat and away from sensible heat with increasing leaf area provides a physical check on the estimated coefficients.

Step 3 — Downwind vegetation response. Monthly anomalies of downwind leaf area are regressed on the moisture and heat advected to the sink, on incoming precipitation, and on the persistence of leaf area from the preceding month. The moisture predictor is the advected quantity delivered from upwind sources rather than the local column water vapour, so that the coefficient measures the vegetation response to transported moisture. Predictors are restricted to quantities exogenous to the downwind vegetation state; the attribution of the delivered conditions to upwind greening is carried by the transport efficiencies from Step 1.

Step 4 — Feedback coefficient. The three components are combined into a source–sink feedback coefficient that measures the change in downwind leaf area produced by a unit change in upwind leaf area. The coefficient separates into a hydrological term, in which the flux response to greening is delivered downwind as precipitation and water vapour and converted into a vegetation response, and a thermal term, in which the sensible heat anomaly is delivered downwind and modifies the vegetation response through temperature. Evaluated over the domain, the coefficient yields the spatial structure of the feedback and identifies the dominant pathway in each regime.

Study domain and period. The analysis covers Europe at the native resolution of the tracking framework (0.5x0.5 degrees), comprising 3,976 vegetated grid cells after sparsely vegetated cells are masked. The framework is run over 20 years, which provides the sample size required for stable monthly per-cell estimates of the transport efficiencies and regression coefficients and spans the range of interannual variability in moisture supply.

A.5. Preliminary Results

In a proof-of-concept study, we tested the self-intensification hypothesis over the coterminous United States (CONUS) using 8 years of FLEXPART data (2003–2010). We found that regions of strong land–atmosphere coupling such as the Midwest and southern US show strong self-intensification of vegetation greening. This project will apply the same methodology to the European domain.

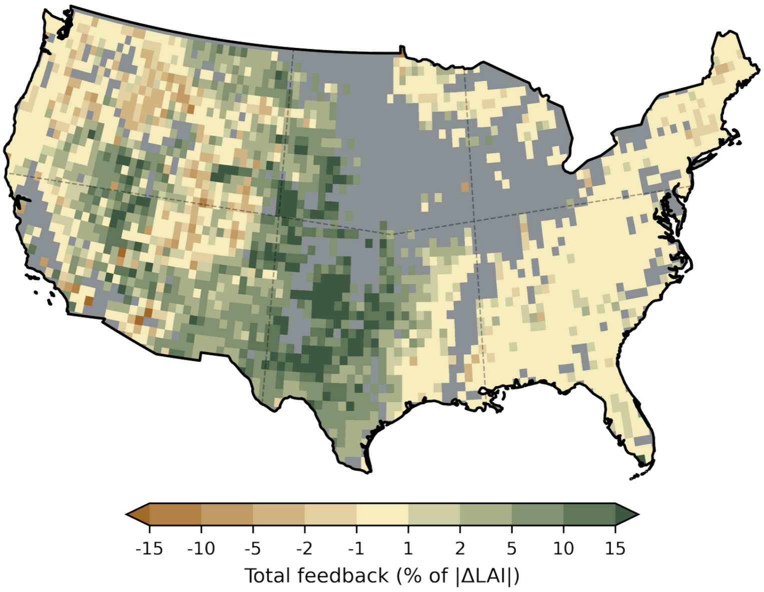

Figure 2: Preliminary Results. *Self-intensification quantified over the growing season using 8-years over the United States.*

B. Technical Plan

B.1. Characteristics of the Code

The moisture and heat tracking is performed with HAMSTER, an open-source framework written in Python that post-processes the trajectory output of the FLEXPART model, itself driven by ERA5. The workflow proceeds in sequential stages: diagnosis of evaporation, precipitation, and sensible heat along trajectories; conversion of the diagnosed output into per-parcel trajectories; attribution of moisture and heat sources for each sink; and bias correction of the attributed sources against observations. The diagnosis stage has been completed under an existing allocation. The resources requested here cover the three remaining stages for the full domain and period: evaluation of trajectories, source attribution, and bias correction.

The workflow is not internally parallelised. The domain is partitioned into subsets of grid cells, and each stage is executed as a Slurm array job in which each task processes one subset independently. The stages are sequential, each depending on the output of the preceding one, and are run one year at a time, with completed output moved to archive before the next year begins.

B.2. Justification of the Resources

The computational cost is estimated per grid cell and year from the core allocation and runtime of the three stages, then scaled to 3,976 grid cells over 20 years. Usage is charged in System Billing Units as SBU = P · N · T, where P = 424074479 / (8090 × 128 × 86400) is the fixed proportionality factor, N is the number of physical cores allocated, and T is the elapsed time in seconds.

Stage	Cores (N)	Runtime per cell-year	SBU per cell-year	20-year total (SBU)
Trajectory conversion	256	2.2 min	158	12,580,700
Attribution	1	3.0 h	51	4,070,600
Bias correction	1	0.75 h	13	1,017,800
Total			**222**	**17,669,100**

Table 1. Estimated SBU cost of the three tracking stages, per grid cell and year and accumulated over 3,976 grid cells and 20 years.

The attribution and bias-correction stages each request a single core, with per-cell-year runtimes of approximately 3 hours and 0.75 hours; their memory footprints fall within the per-core memory of the fractional queue, so each is charged for one core. The trajectory-conversion stage requests a memory allocation (480GB of memory) that exceeds the capacity of a single node and therefore runs on the parallel queue across two nodes, charged for their 256 physical cores, at a per-cell-year runtime of approximately 2.2 minutes. The resulting per-cell-year and full-project costs are summarised in Table 1; the three stages together cost approximately 222 SBU per cell-year, amounting to approximately 17.7 million SBU over 3,976 grid cells and 20 years.

The number of parcels evaluated over the domain varies between years, and operational factors such as occasional job resubmission introduce further variation in consumed resources. A margin is added to accommodate this, giving a request of 18 million SBU for the project year.

Data storage. The tracking input for one year is the global FLEXPART output of approximately 3.6 TB, which is preprocessed to an intermediate dataset and then reduced to the attributed source fields. Processing is performed one year at a time, and completed output is moved to archive before the next year is processed, so transient scratch use remains near 4 TB and does not accumulate across years. The archived output is the attributed source dataset, approximately 600 GB per year, which accumulates to 12 TB over 20 years. A total archive allocation of 15 TB is requested to hold this output together with the retained diagnostic fields and a margin for interannual variation in the volume of tracked data.

References

Duveiller, G., Hooker, J., & Cescatti, A. (2018). The mark of vegetation change on Earth's surface energy balance. Nature Communications 9, 679.

Forzieri, G., Miralles, D. G., Ciais, P., Alkama, R., Ryu, Y., Duveiller, G., Zhang, K., Robertson, E., Kautz, M., Martens, B., Jiang, C., Arneth, A., Georgievski, G., Li, W., Ceccherini, G., Anthoni, P., Lawrence, P., Wiltshire, A., Pongratz, J., Piao, S., Sitch, S., Goll, D. S., Arora, V. K., Lienert, S., Lombardozzi, D., Kato, E., Nabel, J. E. M. S., Tian, H., Friedlingstein, P., & Cescatti, A. (2020). Increased control of vegetation on global terrestrial energy fluxes. Nature Climate Change 10, 356–362.

Keune, J., Schumacher, D. L., & Miralles, D. G. (2022). A unified framework to estimate the origins of atmospheric moisture and heat using Lagrangian models. Geoscientific Model Development 15, 1875–1898.

Koppa, A., Keune, J., Schumacher, D. L., Michaelides, K., Singer, M., Seneviratne, S. I., & Miralles, D. G. (2024). Dryland self-expansion enabled by land–atmosphere feedbacks. Science 385, 967–972.

Stohl, A., Forster, C., Frank, A., Seibert, P., & Wotawa, G. (2005). Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. Atmospheric Chemistry and Physics 5, 2461–2474.

Yu, Y., Notaro, M., Wang, F., Mao, J., Shi, X., & Wei, Y. (2017). Observed positive vegetation–rainfall feedbacks in the Sahel dominated by a moisture recycling mechanism. Nature Communications 8, 1873.

Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., Ciais, P., Sitch, S., Friedlingstein, P., Arneth, A., et al. (2016). Greening of the Earth and its drivers. Nature Climate Change 6, 791–795.