The MMX Toolkit

Broadening the application of ecological science to private- and public-sector conservation efforts

John Schnase and Mark Carroll

Data Science Group / Innovation Lab
NASA Goddard Space Flight Center

Overview

This presentations describes a multi-decadal time series analysis of changing environmental suitability for Cassin's Sparrow

The resulting insights into the species' past response to a changing climate is proving useful to state-level conservation planning

Our approach to retrospective ecological niche modeling was entirely automated and is potentially deployable at an operational scale

We are beginning to assemble the code set used in this research into to what we refer to as the MMX Toolkit

In this talk, we describe the project, our results, and the implications we see for broader use of the technology in conservation practice …

Background

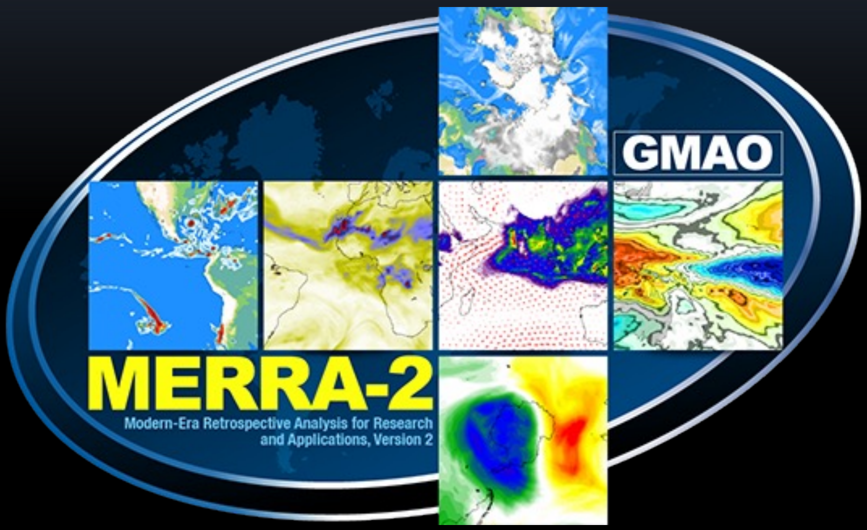
Cassin's Sparrow (*Peucaea cassinii*) was our case study species

- Cassin's Sparrow is a grassland resident of the American Southwest
- Conservation status is poorly understood
- Past responses to a changing climate could help explain uncertainties
- In this study, we combined data from NASA's Modern-Era Reanalysis for Research and Applications, Version 2 (MERRA-2) with field observations spanning the past 40 years to investigate multi-decadal changes in climatic suitability for the species
- Software suite developed for the study resulted in the MMX Toolkit …

Method

The MERRA-2 reanalysis was our sole source of environmental predictors

- Source: Global Modeling and Assimilation Office (GMAO), a configuration of GEOS 5
- Input: hundreds of observation types, millions of observations/day)
- Output: a global temporally and spatially consistent synthesis of hundreds of key climate variables
- Spatial resolution: $1/2^\circ$ latitude $\times$ $2/3^\circ$ longitude $\times$ 42 vertical levels extending through the stratosphere.
- Temporal resolution: 6-hour or less for three-dimensional, full spatial resolution, extending from 1979-Present <== *(Basically, the entire satellite era …)*

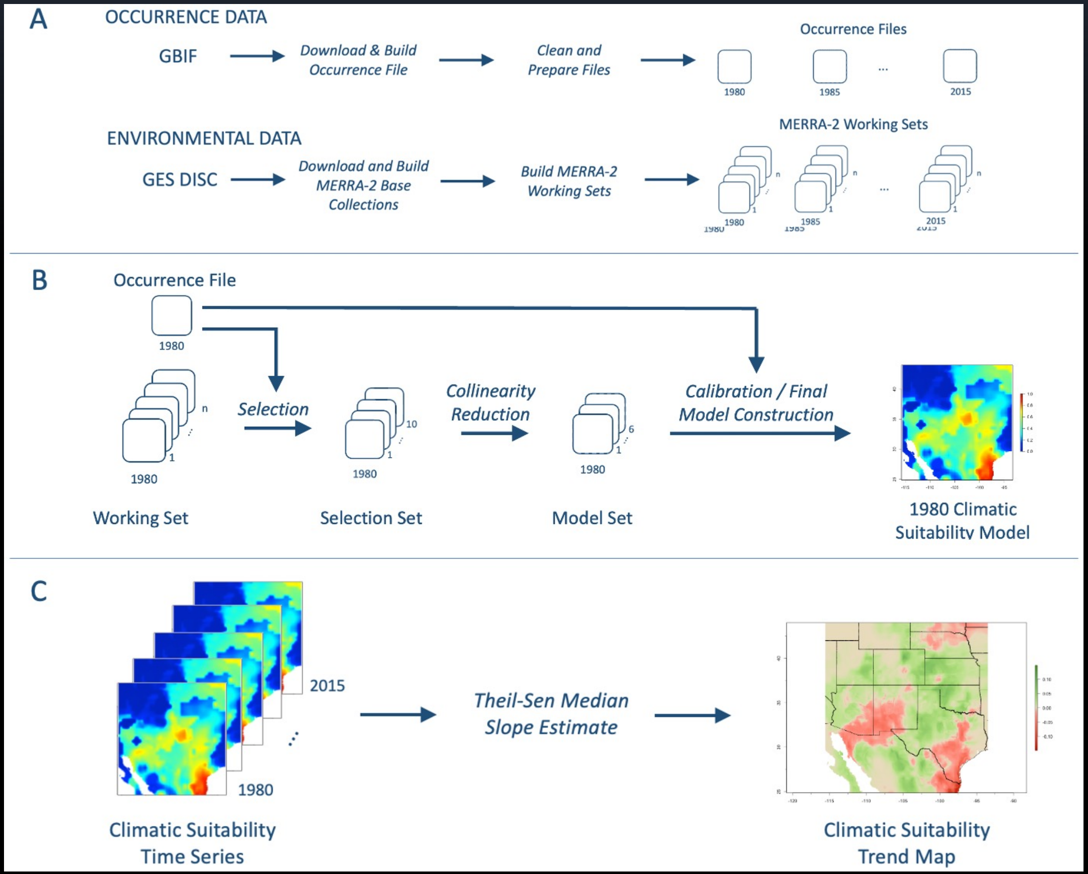


Method

We examined two 40-year time series using MaxEnt

(A) Data assembly:

- The M2 time series used a mix of 30 MERRA-2 ecosystem functional attributes and microclimatic variables
- The MC time series used 19 MERRA-2-derived bioclimatic variables
- Occurrence records from GBIF spanning 1980-2019 in 5-year intervals

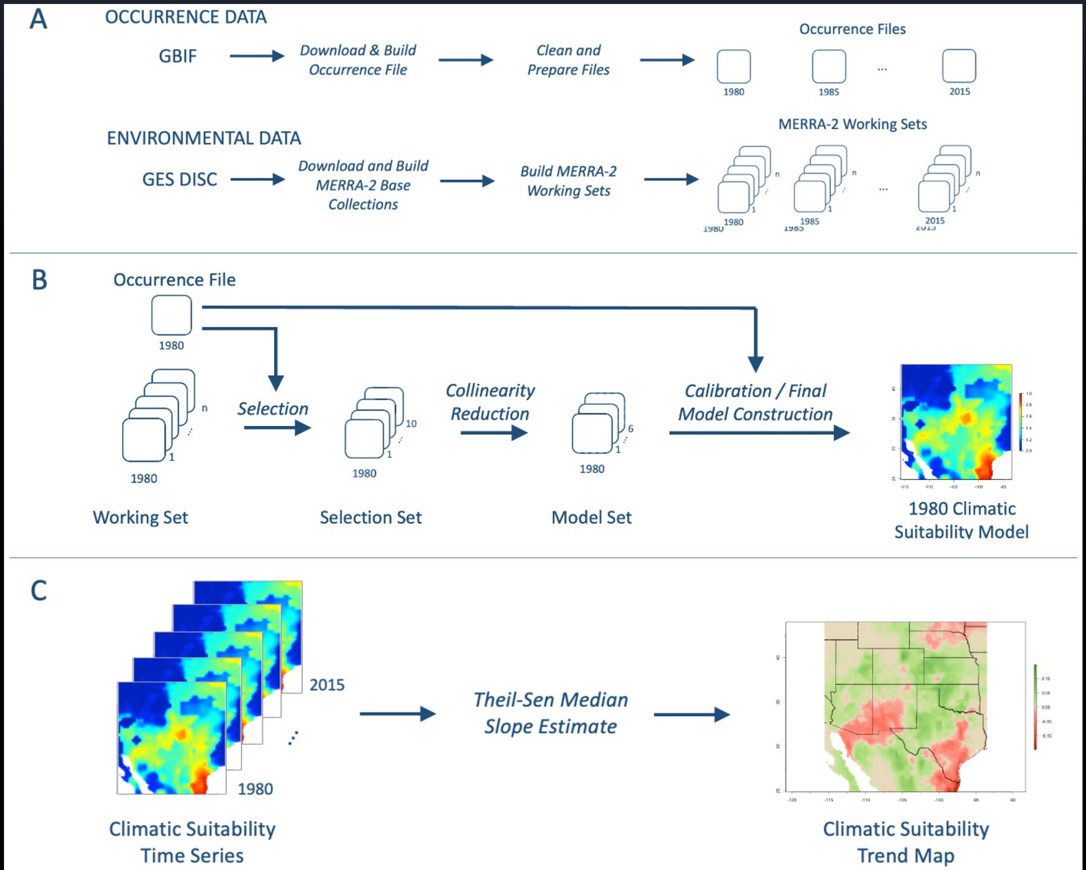


Method

The analysis combined time and variable specificity

(B) Time series construction:

- Used MERRA/Max to do automatic variable selection for each 5-year interval of the time series
- Used VIF analysis to reduce collinearities
- Used ENMeval to tune final models in each 5-year interval of the M2 and MC time series

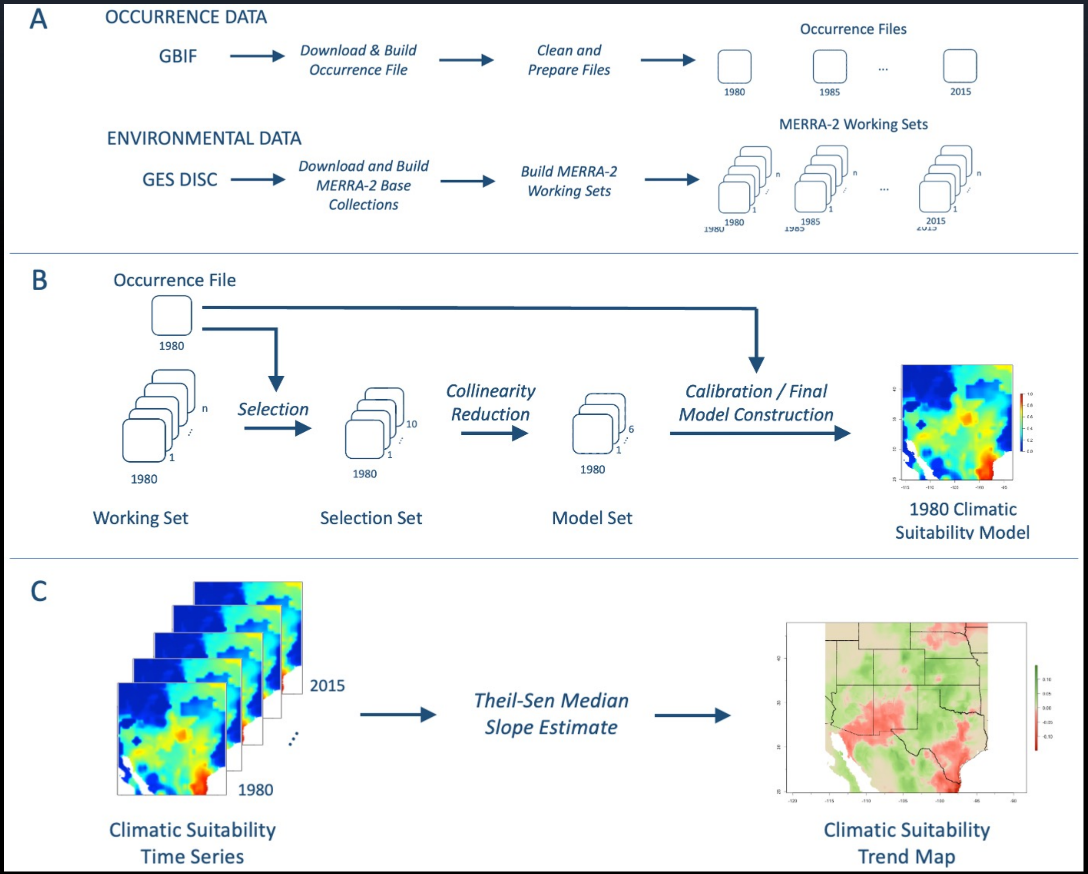


Method

We looked at changes in suitability and environmental drivers

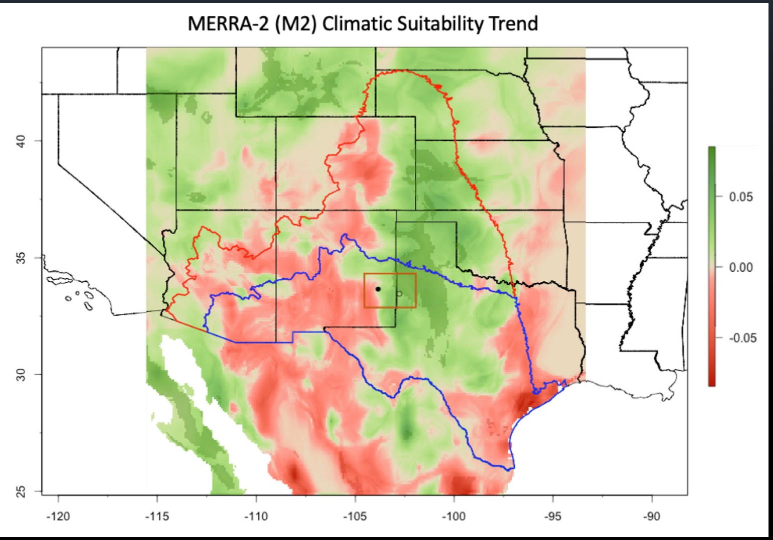
(C) Time series analysis:

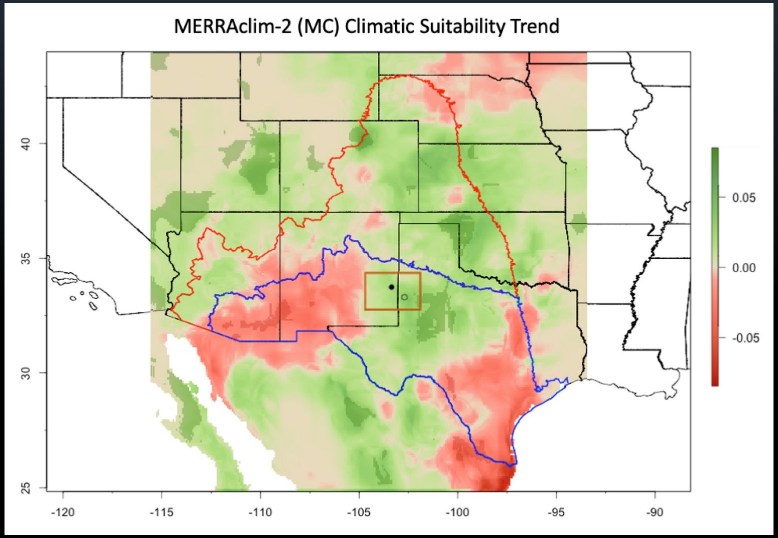
- Used Theil-Sen median slope estimate to determine 40-year trend
- Used weighted centroids to determine 40-shift in suitability
- Did trend analysis on environmental suitability and environmental drivers

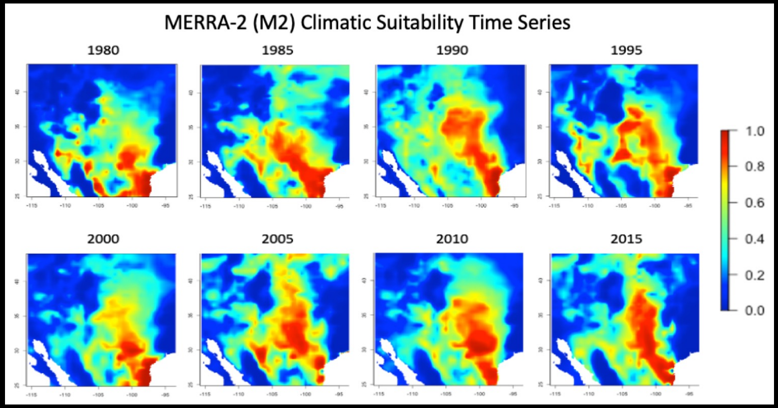


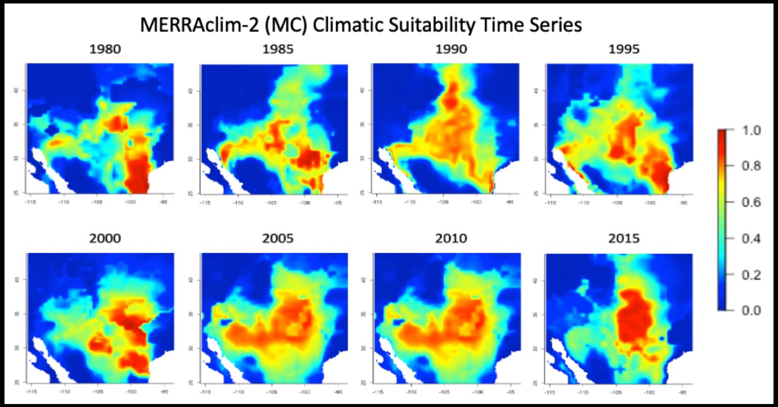
Results

Complex patterns of change seen in M2 and MC time series and trends







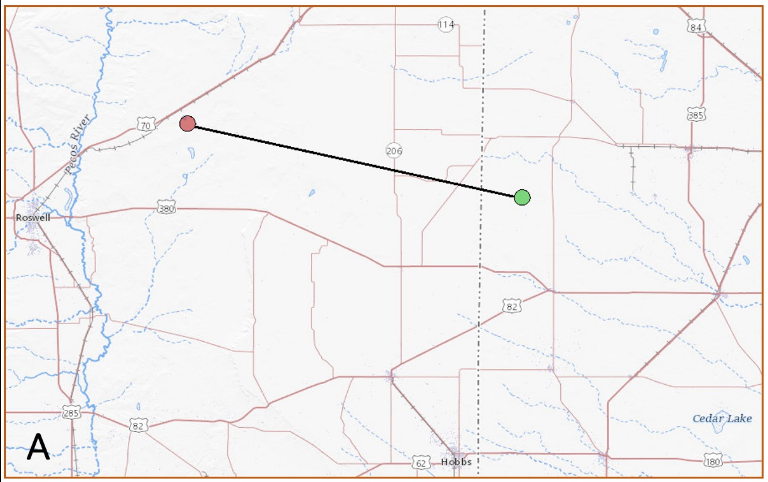


Results

Contrasting patterns of change seen in M2 and MC velocities and areas

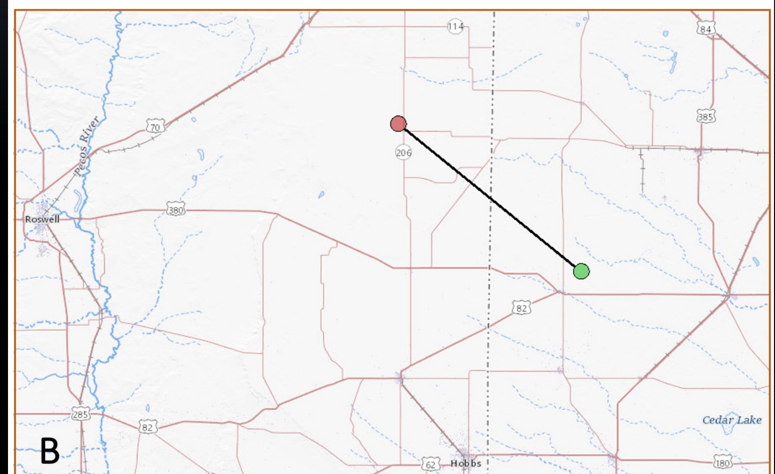
MERRA-2 (M2) Climatic Suitability Centroids

	Latitude	Longitude	Bearing	Distance (km)	Velocity (km/yr)
1980	33.447	-102.916	—	—	—
2015	33.656	-104.012	282° 33' 24''	104.20	2.60



MERRAclim-2 (MC) Climatic Suitability Centroids

	Latitude	Longitude	Bearing	Distance (km)	Velocity (km/yr)
1980	33.256	-102.739	—	—	—
2015	33.679	-102.352	309° 23' 07''	73.82	1.85



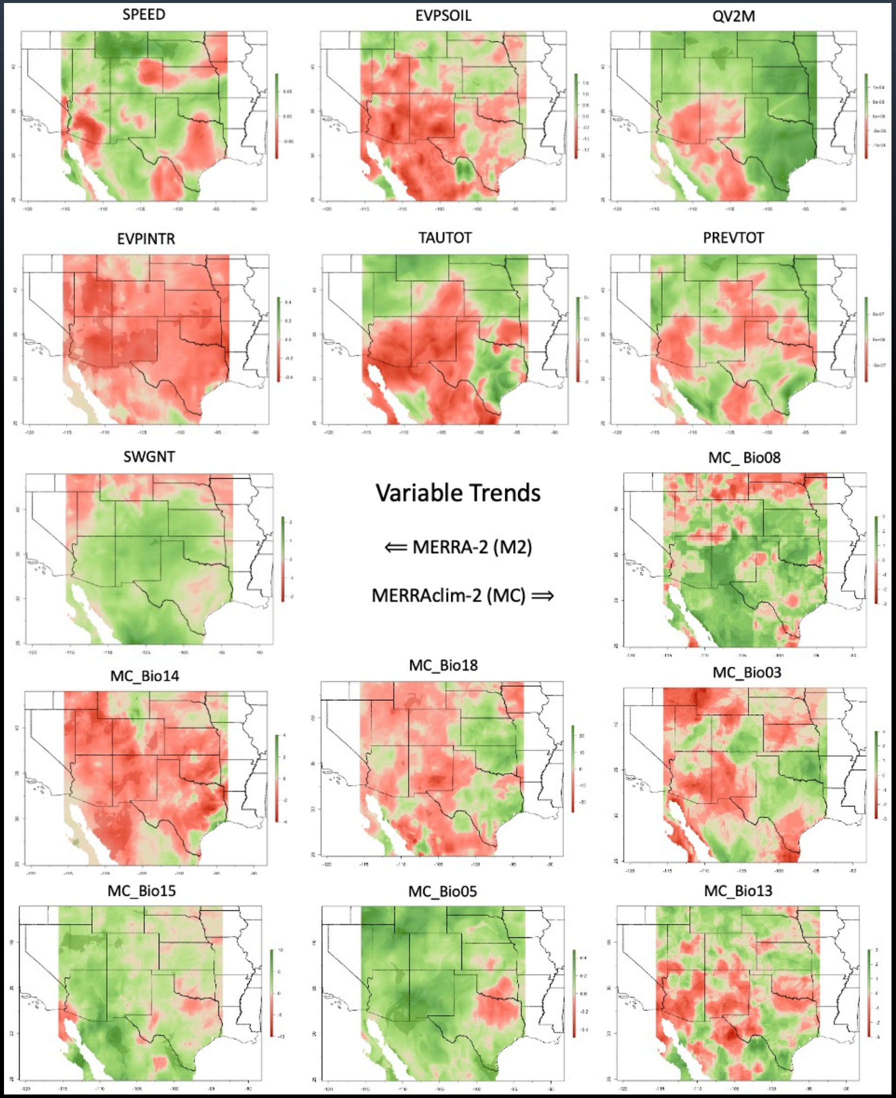
	Overall Trends		Significant Trends			
Time Series	**km² x 10⁵**	**% of area**	**Z**	**km² x 10⁵**	**% of area**	**TS Δ / 5-yr**
MERRA-2 (M2)	20.7	53.0	3.09 ± 0.01	2.41 ± 0.64	6.2 ± 1.6	0.03 ± 0.05
	15.8	40.4	–2.85 ± 0.15	0.19 ± 0.07	0.5 ± 0.2	–0.05 ± 0.01
MERRAclim-2 (MC)	27.3	69.7	3.34 ± 0.08	2.15 ± 0.11	5.5 ± 0.3	0.03 ± 0.01
	0.9	23.9	–2.09 ± 0.01	0.37 ± 0.05	0.9 ± 0.1	–0.06 ± 0.02

Results

Complex trends observed in M2's and MC's top contributing variables

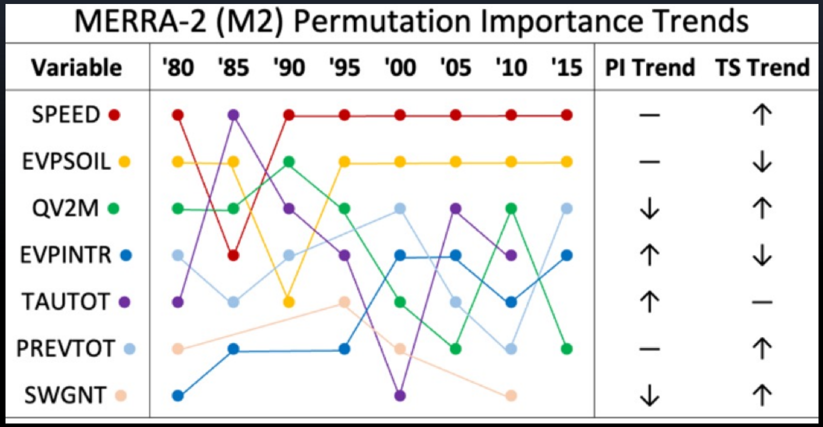
MERRA-2 (M2)	n	PI	Description
SPEED	24	18.3 ± 1.7	Surface wind speed (m/s)
EVPSOIL	24	12.4 ± 1.9	Bare soil evaporation energy flux (W/m2)
QV2M	24	6.2 ± 1.2	2-meter specific humidity (kg/kg)
EVPINTR	20	19.3 ± 1.5	Interception loss energy flux (W/m2)
TAUTOT	17	9.9 ± 1.8	Optical thickness of all clouds
PREVTOT	15	16.4 ± 2.9	Total re-evaporation/sublimation of precipitation ((kg/m2)/s)
SWGNT	13	20.8 ± 2.8	Surface net downward shortwave flux (W/m2)
MERRAclim-2 (MC)	**n**	**PI**	**Description**
MC_Bio08	24	39.6 ± 3.1	Mean temperature of the wettest quarter (°C)
MC_Bio14	19	11.1 ± 1.0	Precipitation of the driest month (mm)
MC_Bio18	18	23.2 ± 2.9	Precipitation of the warmest quarter (mm)
MC_Bio03	16	10.7 ± 2.0	Isothermality [(MC_Bio2/MC_Bio07)*100] (%)
MC_Bio15	15	10.6 ± 1.3	Precipitation seasonality (Coefficient of variation) (%)
MC_Bio05	15	9.2 ± 3.7	Maximum temperature of the warmest month (°C)
MC_Bio13	12	8.1 ± 2.6	Precipitation of the wettest month (mm)

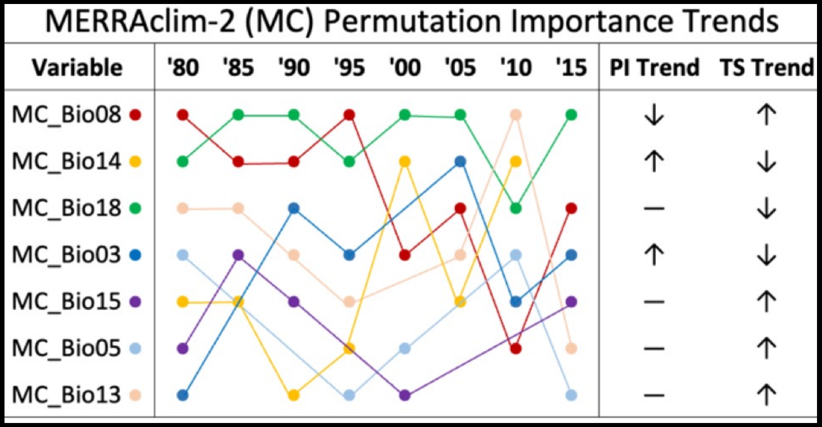
Variable	% of area	Variable	% of area
SPEED	62.5 -29.8	**MC_Bio08**	67.5 25.5
EVPSOIL	34.4 -63.1	**MC_Bio14**	17.2 73.3
QV2m	75.0 -22.3	**MC_Bio18**	36.5 62.1
EVPINTR	5.0 76.7	**MC_Bio03**	44.4 53.4
TAUTOT	46.1 53.4	**MC_Bio15**	80.0 14.8
PREVTOT	57.5 41.7	**MC_Bio05**	85.6 7.9
SWGNT	79.5 20.4	**MC_Bio13**	54.6 43.7

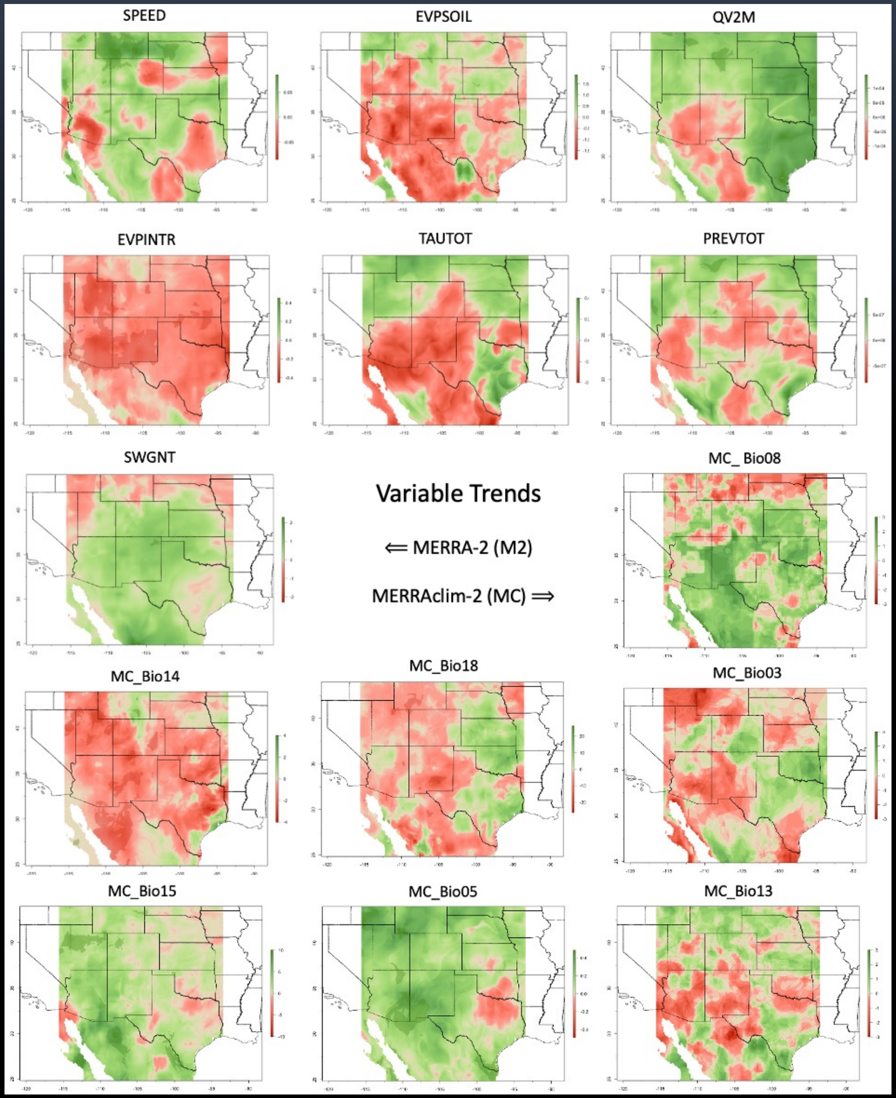


Results

Complex trends observed in the relative contributions of variables







Findings

Complex patterns of changing environmental suitability for Cassin’s Sparrow over the past 40 years

We see that patterns of historical change in environmental suitability can vary widely depending on the spatial extent considered and on whether time series models are driven by bioclimatic variables alone or by variables more aligned with ecosystem functioning. It is likely that demographic and other observational surveys vary greatly across the species' range as well, which could lead to differing conclusions regarding the conservation status of the species.

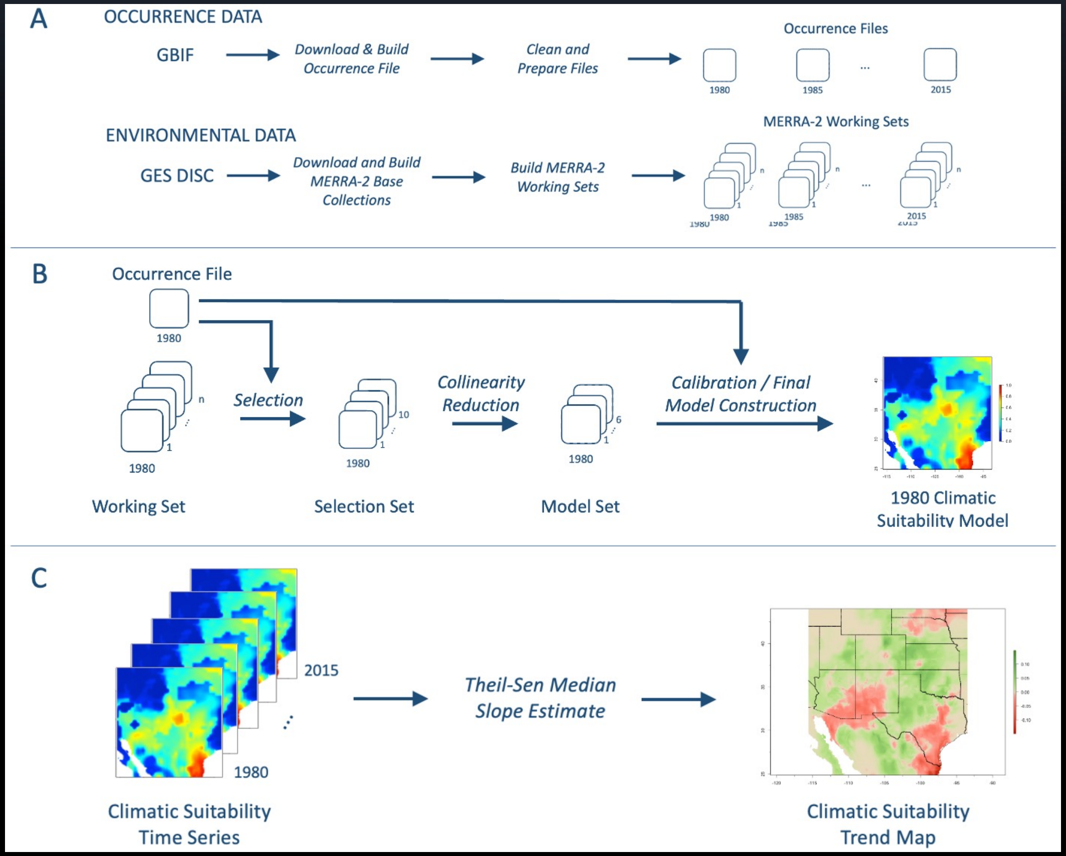
Complex patterns of changing environmental drivers of suitability for Cassin’s Sparrow over the past 40 years

It appears that spatiotemporal attributes of North American Monsoon precipitation, hydrologic conditions of the soil, and surface winds should be considered variables of particular relevance to the Cassin's Sparrow and key to understanding the species' conservation status.

The MMX Toolkit

A prototype software suite to enable retrospective ENM

- The MMX Toolkit bundles the capabilities of NASA's MERRA/Max system with utilities, data, documentation, and other resources to enable historical, ENM-based analysis of the changing climatic suitability of bird species
- Essentially automates the multi-step trend analysis workflow used in this project
- Significantly reduces the time and effort required for retrospective ENM



The MMX Toolkit

Reduces the time and effort required for retrospective ENM

Toolkit operation

- MERRA/Max selection the rate-limiting workflow step
- MERRA/Max's MaxEnt Monte Carlo selection runs are parallelizable
- MERRA/Max operates on filesystem files, therefore infinitely scalable
- We used a set of ten 10-core Debian Linux 9 Stretch VMs
- Other MMX Toolkit software runs on a MacBook Pro connected to remote services

For example …

With 30 and 19 variables respectively in the M2 and MC collections, at 100 samples each performed in triplicate, MERRA/Max selection across the two time series required 7,350 independent bivariate MaxEnt sampling runs and a total computation time of about 90 minutes in our 100-core cluster computer environment.

By contrast, single-processor run time for this computational load would have been about four days.

In theory, the approach described here would require less than one minute in a fully provisioned 7,350-core cluster computing environment, which makes an otherwise intractable manual analysis possible.

Commercial cloud computing makes a service such as this affordable and accessible …

The MMX Toolkit

In active development, please feel free to follow our progress …

- Schnase J.L., Carroll M.L., et al. Toward a Monte Carlo approach to selecting climate variables in MaxEnt. PLOS ONE. 2021;16: e0237208. doi:10.1371/journal.pone.0237208
- Schnase, J.L. and Carroll, M.L. Automatic variable selection in ecological niche modeling. PLOS ONE. 2022;17: e0257502. doi:10.1371/journal.pone.0257502.
- Schnase, J.L. and Carroll, M.L. 2023. The MMX Toolkit: High performance, reanalysis-based climatic suitability modeling to advance avian conservation. 2023 Big Data in Space Conference (BiDS’23). In review.
- Schnase, J.L., Carroll, M.L., Montesano, P.M., and Seamster, V.A. Complex changes in climatic suitability for the Cassin’s Sparrow (*Peucaea cassinii*) revealed by retrospective ecological niche modeling. In prep.

For more information, please contact
John.L.Schnase@NASA.gov
Code 606.3 – Data Science Group / Innovation Lab
NASA Goddard Space Flight Center