Building an AI-Ready Camera–METAR Dataset from Publicly Available Data for Visibility and Cloud-Ceiling Estimation

UND AEROSPACE
UNIVERSITY OF NORTH DAKOTA
Atmospheric Sciences

Jenna Post[1], Marwa Majdi[1] and David Delene[1]

[1]Department of Atmospheric Sciences, University of North Dakota, United States of America

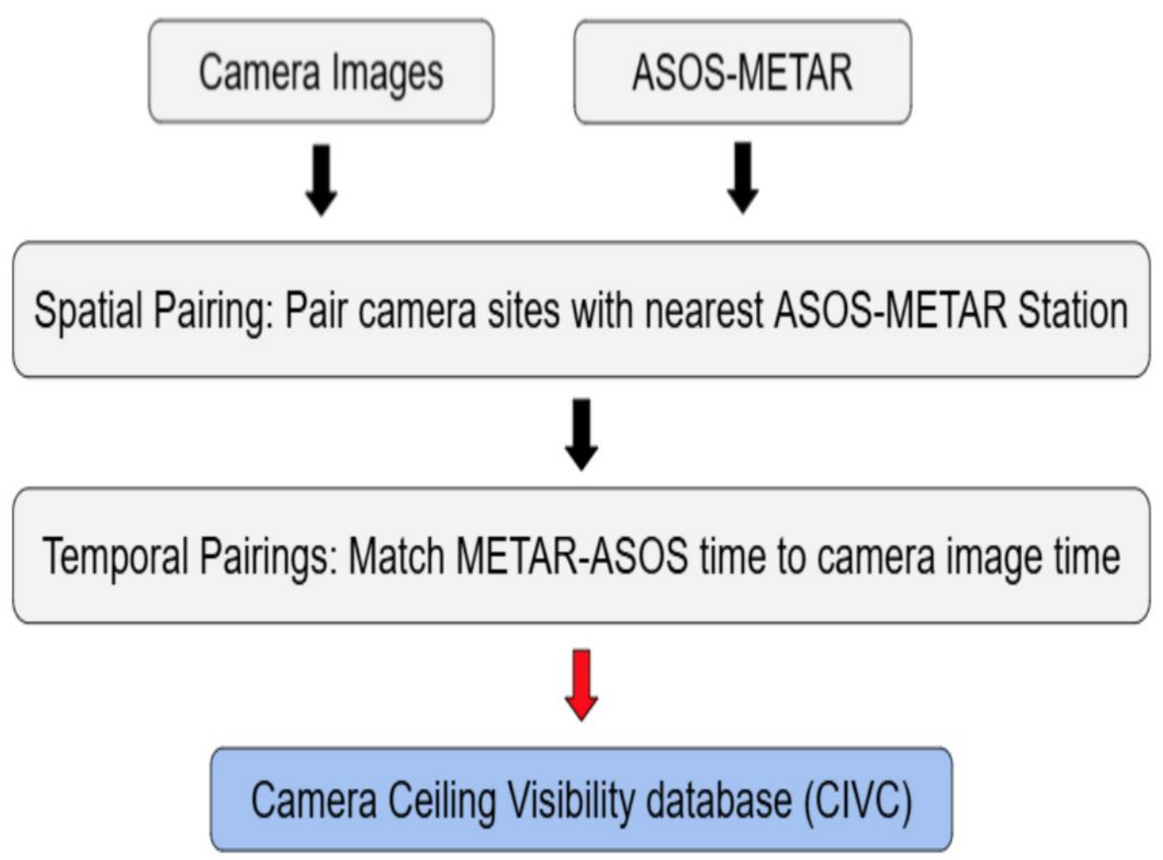
Motivation and Objective

- Visibility and cloud ceiling are critical for aviation and UAS operations.
- METAR/ASOS observations are spatially limited.
- Publicly available fixed-camera networks provide continuous visual information on sky, horizon, and scene conditions.
- AI-based estimation requires a structured, reproducible, and well-labeled dataset.

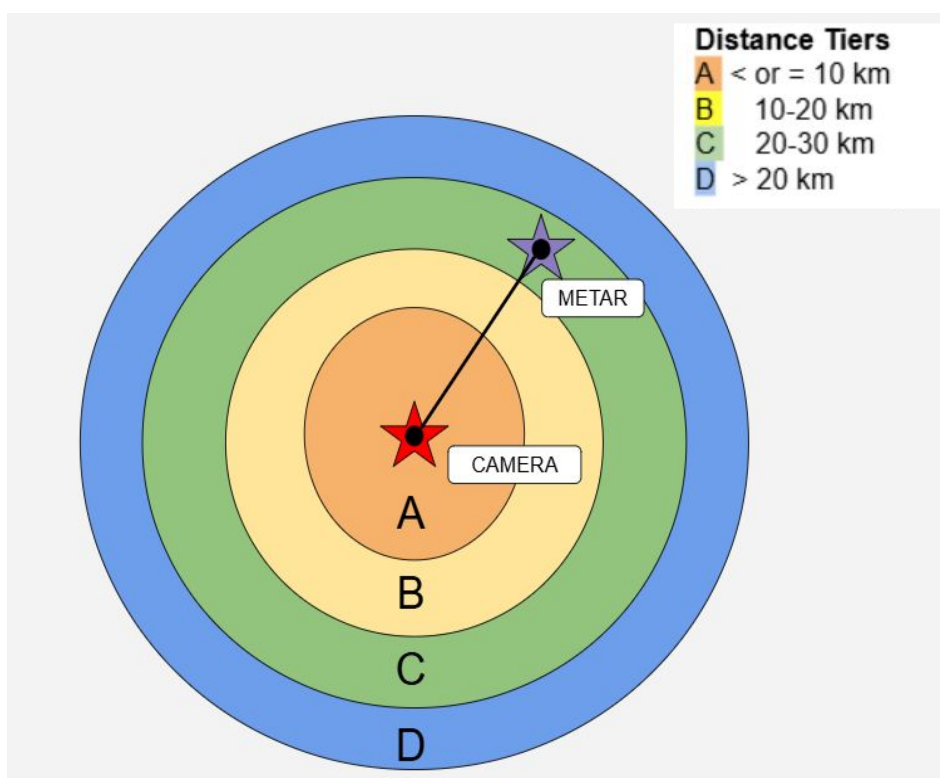
Objective: Develop the **Camera Image and Visibility–Ceiling (CIVC)** dataset from publicly available camera and METAR data for AI-based visibility and cloud-ceiling determination.

Data and Methodology

CIVC Dataset Development Workflow



Distance Tiers for Spatial Pairing



- Public camera images are matched with nearby METAR/ASOS observations through spatial and temporal pairing, followed by quality control.
- Camera sites are grouped by distance from nearby METAR stations to support spatial pairing. These distance tiers help evaluate pairing proximity and matching quality.
- Temporal pairing matches each image to the nearest METAR/ASOS report in time.

Results

1. Representative Camera Images Over North Dakota

Clifford Hall (UND Camera)

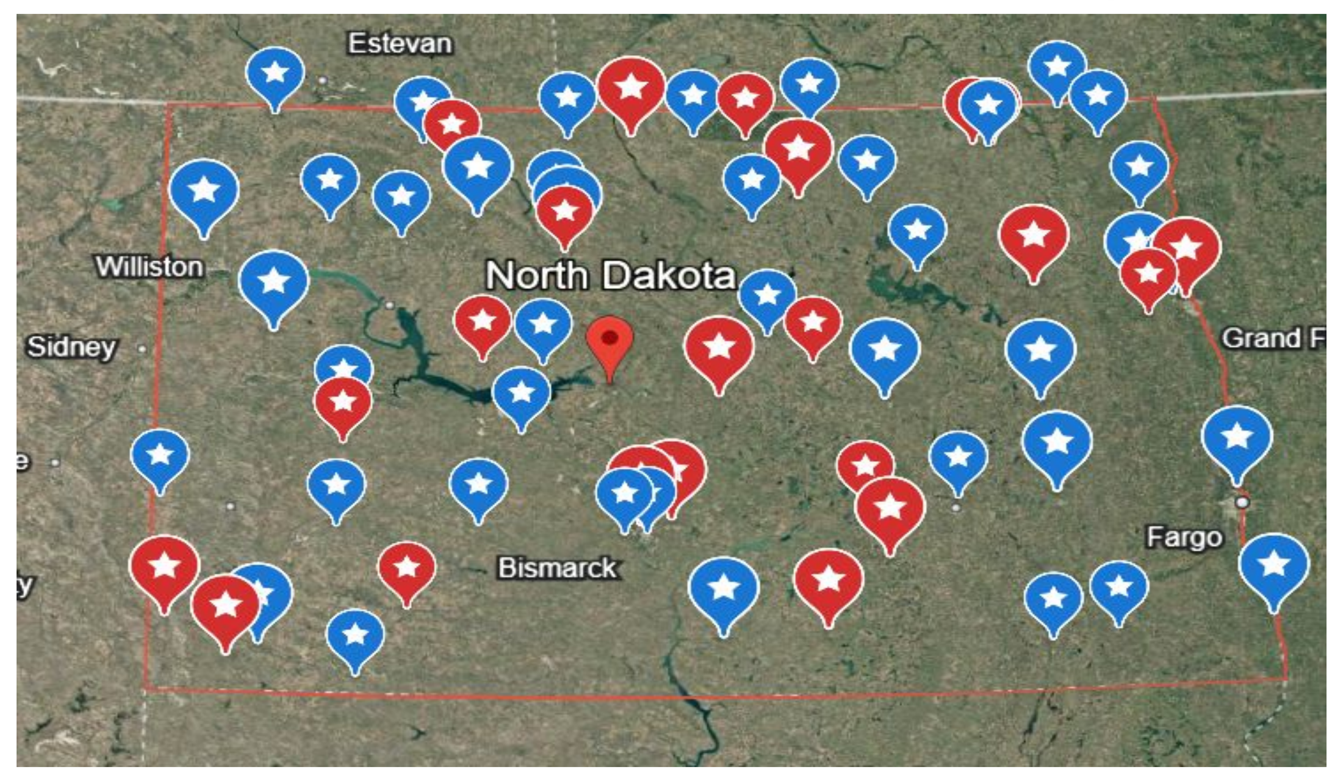
Dunseith, ND (NDDOT Camera)

Bowman, ND (NDDOT Camera)

- North Dakota camera scenes are collected from public networks.
- These images provide sky, horizon, and foreground information relevant to visibility and ceiling.

2. North Dakota Camera Sites and METAR Stations

North Dakota camera sites and METAR stations used for spatial pairing



Camera Sites (22)

METAR/ASOS Stations (39)

- North Dakota camera sites and METAR stations are mapped.
- The site distribution supports statewide spatial pairing.

3. CIVC Database Structure

	Camera						Metar-ASOS				Pairing			Data	Quality Check	
State	network	id	lon	lat	time	path	id	lon	lat	time	dis	tiers	dt	metar vis+sky	pair flag	quality pair

- The CIVCdatabase stores one row per matched camera image-METAR record.
- Each record includes image metadata, station metadata, visibility, cloud-ceiling, and pairing flags.

Conclusions

- A reproducible automatic workflow is developed to build the CIVC database from camera images and METAR observations.
- Camera sites are paired with nearby METAR stations over North Dakota in space and time using standardized rules.
- The automatic workflow integrates real-time image collection, timestamp standardization, spatial pairing, temporal pairing, and label extraction.
- The CIVC database provides a foundation for future AI-based visibility and cloud-ceiling determination.

Future Work

- Expand the dataset to additional states and public camera networks.
- Refine pairing-quality and quality-control metrics.
- Train and evaluate machine-learning models using the prepared dataset.

Acknowledgments

This work was supported by the University of North Dakota Aerospace Seed Fund. The authors also acknowledge publicly available camera networks and the Iowa Environmental Mesonet archive for providing data used in this study.