P4.1

TOOLS FOR DISPLAYING THE ROBUSTNESS AND DENSITY OF AN OBSERVING NETWORK OVER TIME

Derek S. Arndt*

1. METADATA PHILOSOPHY

Metadata are a vital component of any network's dataset. They bring richness and integrity to data, just as careful quality assurance techniques ensure fidelity. Metadata are most effective when delivered with – or, ideally, in advance of – the data itself.

Integration of metadata into the data stream is a three-step challenge. First, metadata must be identified, which is perhaps the most difficult of the three steps. Second, metadata must be harvested in a consistent and prudent manner. Third, metadata must be conveyed in a way that is useful, understandable and timely for the data client.

2. INTRODUCTION

With the continued explosion in the use and speed of the world-wide web, the nature of climate data transactions is evolving. Data exchanges that used to take at least one phone call and several days in the mail can now be performed immediately and on a self-serve basis. This automated transaction is advantageous to the data client and the data provider. The client can access the data at any time, and the provider is spared the often rote work of retrieving, compiling and transmitting packages of information.

However, the automated exchange model also contains a few drawbacks. First, the client must know exactly what he needs in the way of data, and cannot rely on the expertise of the provider. Second, the client is often unaware of (and even unaware of the importance of) of crucial metadata associated with his data selection.

In order to help address this second problem, the Oklahoma Climatological Survey (OCS) is developing several tools for pre-assessing the robustness and density of network data. The goal for these tools is to provide information that leads to more intelligent data selection in a "self-serve" data environment.

Corresponding author address: Derek S. Arndt, Oklahoma Climatological Survey, 100 East Boyd, Suite 1210, Norman, OK 73019-1012. e-mail: darndt@ou.edu

3. METADATA TOOLS

The OCS metadata indicators were first implemented and tested on data from the U.S. Cooperative Observer (COOP) Network, but their basic tenets are applicable to any geospatial meteorological dataset. They are designed to assist the data client with decisions regarding the appropriateness and availability of a dataset before he commits to its use. They are available for the entire state or for any of its nine climate divisions (Fig. 1).

Simplicity of presentation was a key goal in their design. The concept of metadata is foreign to many and can be regarded as an unnecessary layer of information to a client primarily interested in acquiring weather or climate data to solve a problem. So, tables were avoided in deference to more visual means of communicating information.

3.1 Number of Stations in the Network

The simplest display of metadata available to the data client is a time series of the number of available stations for the region of interest (Fig. 2). The display provides historical context for the client's period of interest by providing the entire period of record for comparison.

3.2 Maps of Areal Coverage

Oklahoma was settled in small pieces of area in small amounts of time. As a result, the evolution of the COOP network within the state was also spotty and specific to developing regions. Often the total number of stations isn't enough to

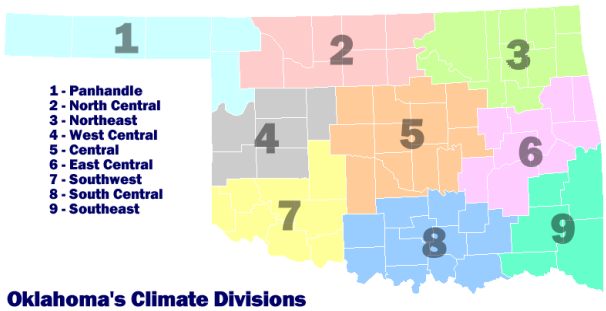

Fig. 1. Oklahoma's Climate Divisions.

demonstrate the health of the COOP network at a given time for a given place. In order for an OCS data client to see the health of the network in his time and region of interest, a series of maps are available for each month since January 1893 (Fig. 3).

The evolution of the state's settlement is apparent in Figure 3. In 1896, just a few years after the opening of the Unassigned Lands in central Oklahoma (April 1889) and the Cherokee Strip in north-central Oklahoma (September 1893), the majority of the state's functioning COOP stations are located in these two regions. At statehood (November 1907), large portions of southeast and northwest Oklahoma contained very sparse populations, and thus very few weather observers.

3.3 Time-series of areal coverage

In order to further compare the robustness of the network over time, a third metadata graphic is displayed to the data client (Fig. 4). The bottom axis represents time, while the vertical axis displays area – specifically, the areas of CD5 within certain distances of a COOP station. For example, the darkest green represents the area within 6 km of a coop station, and the light cream color represents the area of the climate division that was more than 30 km from a coop station.

During the last century, the dark green area has expanded, while the light cream area has decreased. This improvement in station coverage represents two trends: (1) the number of stations has increased, and (2) the spatial distribution of those stations has improved. In fact, the latter is especially true for the last 50 years, because the number of stations has actually decreased slightly since the mid-1950s.

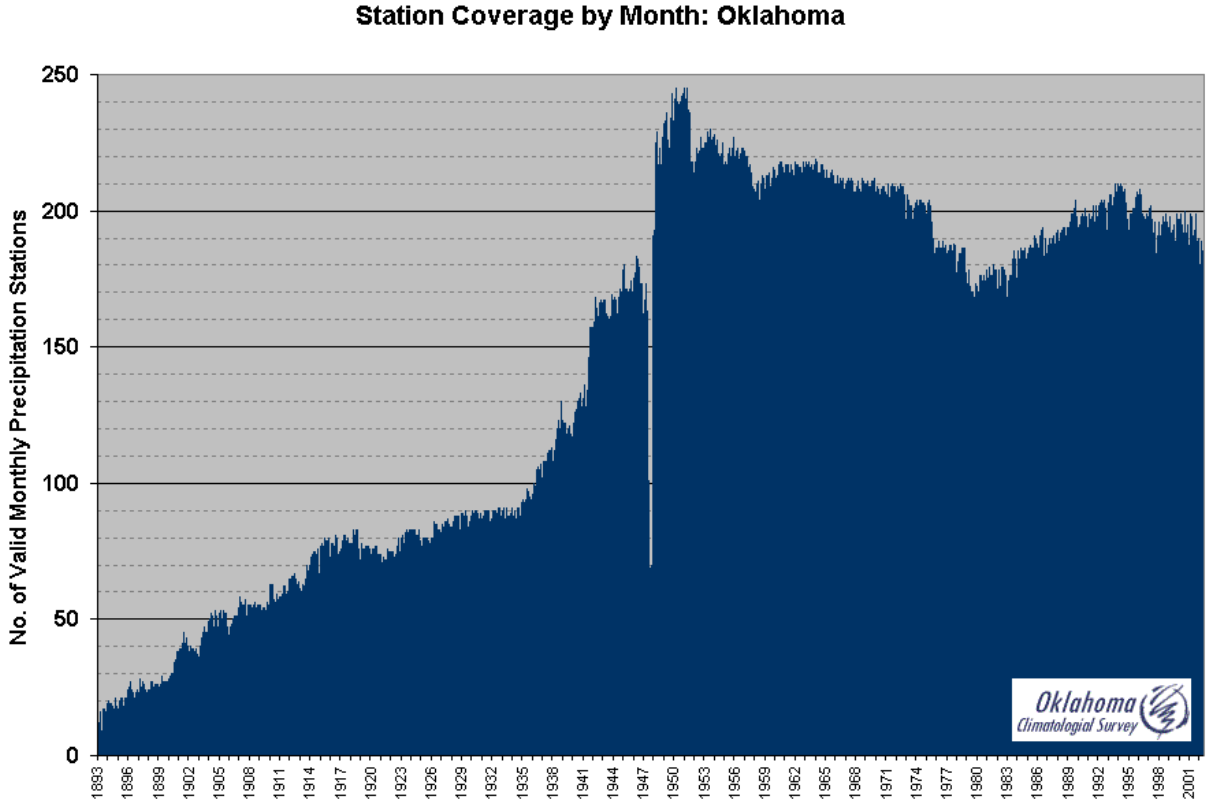

Fig. 2. The number of valid precipitation stations in Oklahoma for each month from 1893-2002.

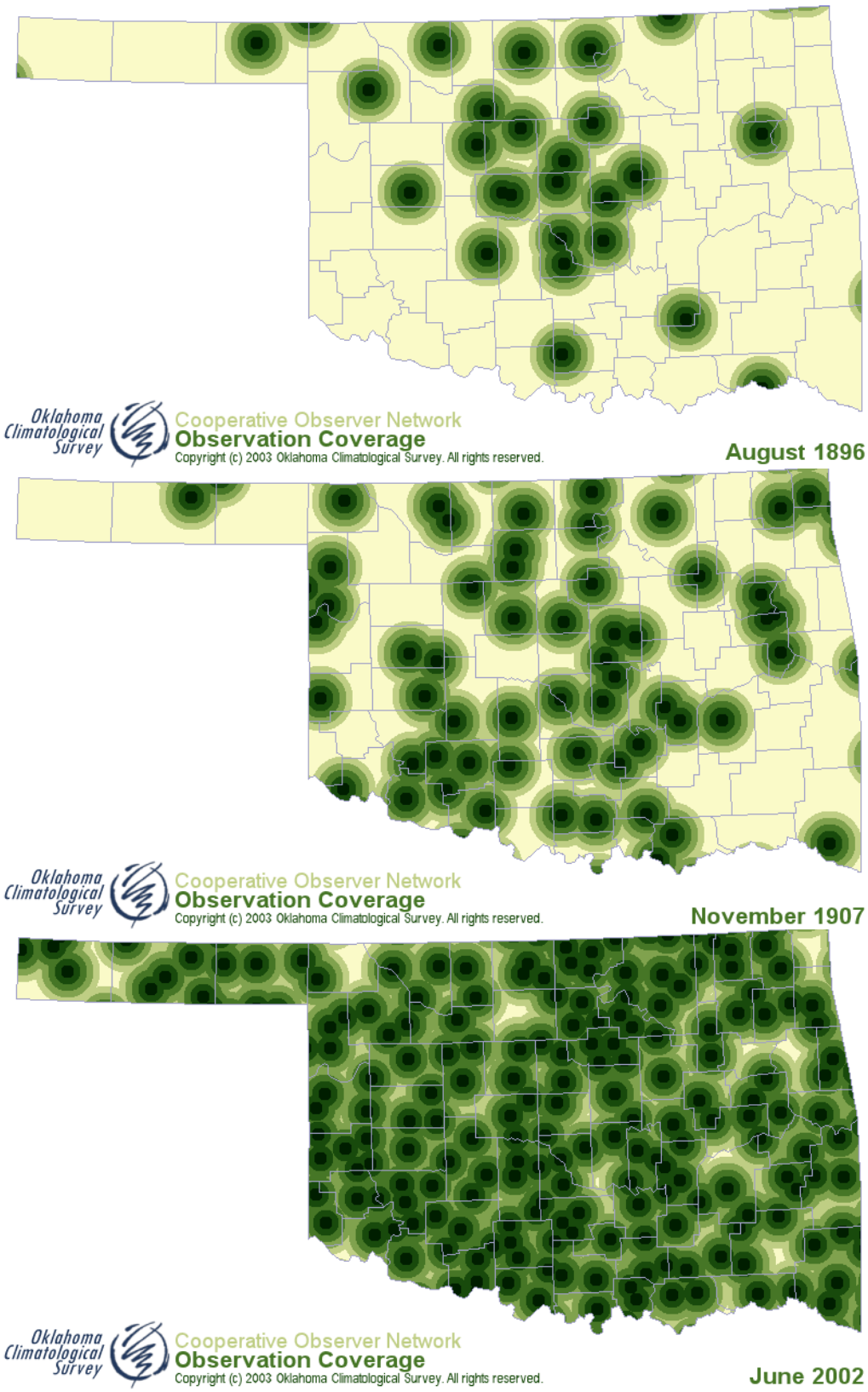

Fig. 3. Spatial distribution of stations within Oklahoma in (a) August 1896, (b) November 1907, and (c) June 2002.

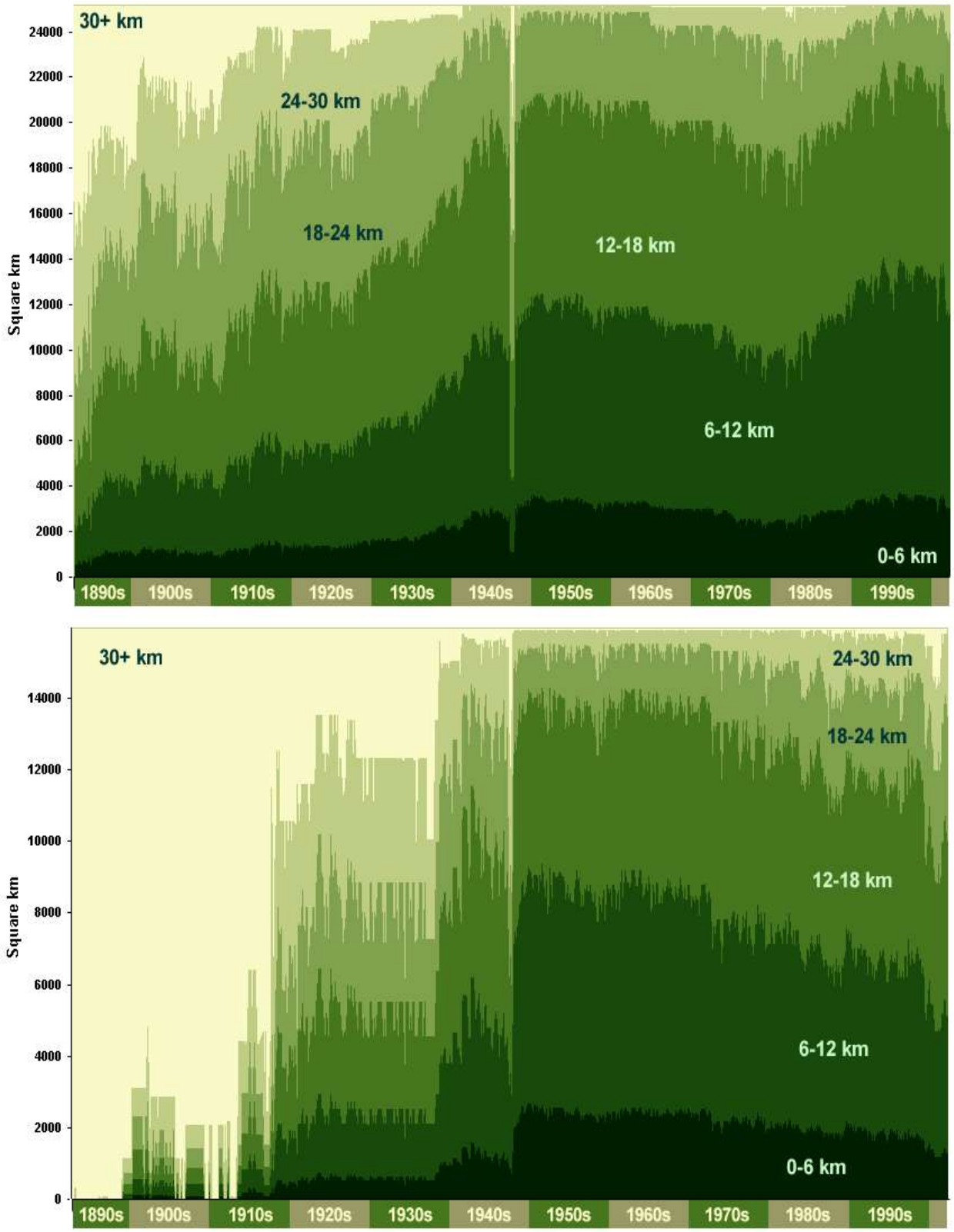

Fig. 4. Amount of area within certain distance of the nearest COOP station for two Oklahoma Climate Divisions: (a) CD5, Central; and (b) CD9, Southeast.