

CPAD

California Protected
Areas Database

CPAD PROJECT WORKING PAPER

MAP BASE LAYERS FOR CPAD: Assessment and Recommendations

Prepared by

Larry Orman
Project Director

March 20, 2013

CONTENTS

Summary

I. Introduction

II. Data Used to Align CPAD Holdings

III. Data Use to Frame CPAD Geometry

IV. Data Used In Relation to Base Layers



www.greeninfo.org

SUMMARY

The California Protected Areas Database (CPAD) is a statewide inventory of parks and other protected open lands owned by agencies and organizations for the purpose of maintaining these open space uses (see www.CALands.org for more information). In late 2012, GreenInfo Network, the originator of CPAD, began a two year project to improve this data set, with funding support from the California Strategic Growth Council through the USGS Gap Analysis Program. This report is one of a series being issued through this project.

This memo assesses a core question for CPAD: what base data should drive the geometric accuracy of its holdings data? The following are its primary recommendations:

1. Continue the default use of assessor parcel data as the primary base layer for CPAD data, in order to ensure overall consistency with emerging state efforts on parcel data.
2. Provide selective options for alternative base alignment data, particularly for broad areas of federal land holdings based on PLSS, and for individual holdings where survey boundaries or aerial imagery indicate significant improvement in locational accuracy over parcel data. Continue to code CPAD as to alignment source used.
3. Use the official California county boundary file for defining county lines. For coastal shorelines, seek to improve the accuracy of this data and provide the file to the State of California for incorporation into the official county file and for general use.
4. Continue to use California Teale Albers as the projection for CPAD data.

I. INTRODUCTION

The creation of effective GIS data requires the use of base data – geographic information used as location reference and correlation for particular thematic data. Typically, base data layers are authoritative (widely accepted as valid) and change rarely, allowing them to be used for data that is itself expected to endure.

This memo assesses a core question for CPAD: what base data should drive the geometric accuracy of its holdings data?

At the outset, it is important to note that CPAD data is not intended to be primary data for defining the boundaries of property ownerships – agencies and organizations owning the holdings in CPAD must be turned to for final determinations of the specific location of their lands.

CPAD as an overall inventory is vital, however, for more general assessments across multiple agency ownerships. CPAD covers lands of almost 1,000 agencies must therefore harmonize alignments among all these holdings in a reasonable manner.

There are three main types of base geographic information that are relevant to CPAD alignment:

1. Data that can be used to specifically define the boundaries in CPAD (e.g., parcel ownership boundaries are used extensively in CPAD)
2. Data that is used to frame or organize CPAD data (e.g., county boundaries or coastlines) across physical or administrative boundaries.
3. Data conventions or standards that shape how data layers are aligned to each other (projections, etc.)

The scale at which any of the above data is authoritative is worth noting – for example, aligning data to a widely used coastline data set created at 1:1,000,000 scale when that data is expected to be used at 1:24,000 or greater would introduce considerable error. CPAD was originally developed at 1:24,000 scale but as parcel data has become more available is now at much larger scales, down to 1:5,000 or better, depending on the base/source data scale.

II. DATA USED TO ALIGN CPAD HOLDINGS

CPAD contains lands that are owned by agencies and other organizations. These ownerships are almost always defined in deeds and other property records and follow legally enforceable boundaries. Using ownership information, therefore, is the default approach employed in CPAD data development.

However, no one system of property boundaries in GIS is truly authoritative – all have weaknesses of varying types and for very specific dispute resolution, custom surveys are often required (but are very expensive). The choice for CPAD, therefore, is one of balancing pros and cons of all options and choosing the most accurate, yet effective, approach.

These sources of base GIS data are used to align CPAD:

- **Source agency GIS records:** CPAD contains holdings of more than 900 public agencies and other organizations. Wherever possible, CPAD practice is to rely on this data to define where ownerships are located, with adjustments as noted below – however, at present, it is feasible to only secure ownership data from the top 100 land owning agencies (comprising 99% of acres).
- **Assessor Parcels:** Each county in California now has a digital (GIS) database of its ownership boundaries and related attribute information. Increasingly, this data is used as a reference layer for many types of GIS data and CPAD has been mainly aligned to these parcel layers, as they allow for more standardized resolution of differences between agencies records. GreenInfo, however, is only able to access data it acquires directly from each county, rather than a statewide data layer.
- **PLSS:** The Public Land Survey System maintained by Bureau of Land Management can be used where parcels are known to have serious accuracy issues (though this is difficult to determine) – and it is somewhat used by federal agencies to define alignments of their holding (see later discussion of issues here).
- **Aerial Imagery:** The availability of high resolution aerial imagery for California enables some boundaries to be referenced directly to such an aerial image. The increasing availability of parcel data helps limit the need for this, but in some cases, imagery is needed to separate developed areas from park space.
- **Reference Layers:** The State’s official county/coastline data and city boundary files are helpful adjuncts for defining alignment (though the city file is not always completely up to date).

DISCUSSION

Source Data Overview

There are 980 agencies and organizations that own the holdings recorded in CPAD. Of these, the top 50 own 98+ percent of all protected land in the state – most of these agencies have GIS data records of their holdings available. The next 50 agencies bring the acreage total to 99% of all lands and many of these have digital data defining their holdings. The remaining 800+ agencies have small acreages but account for 65 percent of the “parks” (SuperUnits) in CPAD, particularly cities with many urban parks. Data availability for these latter agencies varies greatly, with many cities not having parks data easily available in GIS format.

Larger agencies with GIS data usually base it on the following underlying geometries:

1. Parcels (mostly cities and other metropolitan/local agencies now that parcel data covers all counties in the state); CPAD directly incorporates holdings data from agencies who map by parcel (with adjustment of attributes) – this data generally needs no boundary edits.
2. PLSS lines (some federal agencies); For federal agencies using PLSS, CPAD adjusts these holdings to parcels, except in the southern California desert where BLM owns most land and budget has not existed for parcel adjustments (see also below re accuracy of parcel data).
3. Survey records converted to GIS; For agencies using survey-grade GIS data that differs from parcels, engagement is sought with these agencies where possible to work out any differences. These usually involve only individual parcels where there are significant discrepancies from parcel lines – in those cases, CPAD often accepts the certified survey data lines. This process is still evolving, however, and subject to change. In practice, survey data outside of assessor parcels is very challenging to use consistently in CPAD, outside of occasional exceptions.
4. Streets, air photos or other non-parcel/survey source data: This data is usually tightly integrated with parcel data. Where an agency’s CPAD holdings are not aligned to parcels or PLSS, it is usually tested against air photos and streets to ensure accurate alignment.

Pros and Cons of Assessor Parcels

All California counties now have digital/GIS parcel data. Access to and in some cases the quality of this data varies, but some form of parcel data has now become the base for CPAD in all counties (except Sierra, where there have been issues about accuracy).

A separate CPAD report covers the degree of fit between parcels and CPAD, but in general, two thirds or more of CPAD holdings are now aligned to county parcels (a significant acreage of holdings in the southern California desert is more closely aligned to the Public Land Survey System (PLSS) as they are BLM lands).

Use of assessor parcels as a base layer for CPAD has been the default approach for the past five years, marking when parcel GIS data started to become available. In general, CPAD seeks to be as geometrically accurate as possible – whereas early CPAD work was based on digitizing from 1:24,000 scale base maps (USGS quads), current practice often results in 1:5,000 or better scale accuracy (usually 2-10' in most urban/developed areas, but ranging from 100-500' accuracy in some rural areas – for example, Humboldt County caveats inaccuracy of 50' in urban areas and up to 400' in rural areas).

Parcels offer great advantages as the base for CPAD geometry:

- Counties have invested great resources in making accurate determinations of property boundaries – few other agencies have the capacity to improve on these geometries.
- Parcels provide consistency in decision making about boundaries when integrating the holdings of dozens or more agencies in any county – if the State is able to improve its overall approach to parcel data, this will further build consistency.
- California has established efforts to create a statewide parcel layer and have that be a core framework layer for a wide range of GIS-based activities; a related project to this CPAD work is underway at UC Davis to further develop a statewide framework, and the California GIO has recently acquired a statewide commercial parcel database for interim use by state agencies. Previous work on parcels for California was presented in June 2004 in the Digital Lands Record Information (DLRI) Report from the Calif. Mapping Coordination Committee
- In addition to zoning classifications and many general plan designations, many emerging planning tools for land use, conservation and other topics are being built on parcel databases (for example, the new “footprint” models of land use and transportation being developed in major metropolitan areas). Having different geometries limits the value of these regional, state and local planning tools.
- Federal GIS strategy and practice is strongly focused on the goal of a unified cadastral (property) database for the entire country – while a long way off, the broad impetus is for property based geometries to be aligned to cadastral data, with debates/issues regarding PLSS and local or state government parcel

approaches (in some states PLSS and parcels align well, in others not – California tends more toward not).

Parcels, however, are not without issues. Particularly in rural areas, as noted, parcel lines may not be highly accurate, as the exact location of properties in these areas may not be of priority concern to hard-pressed assessor budgets (compared to urban areas where denser populations require high levels of accuracy). However, it is unclear that any other system provides better parcel line accuracy, beyond site by site field surveys.

Access to most-current parcel data can also be a challenge, due to pricing and other restrictions by individual counties (despite legal rulings favoring widespread low-cost access to parcel geometry files).

Finally, parcel data is often highly inconsistent, incomplete and even just wrong when it comes to identifying owners of publicly-owned open space lands. Because these lands are generally not taxable, there is little incentive for counties to have good owner data, with the result being that parcel data cannot always be used effectively to determine if a parcel of land is public open space or which agency might actually own it.

PLSS – Federal Policy, But Limited Usefulness in California

Created in the late 1700s, this system relies on rectangular grids adjusted for particular circumstances to be the framing elements for ownership boundaries – PLSS is most known for its township-range-section geometry which works from 640 acre cells down to small plots split through quartering each size increment (rarely below 10 acres, however). In principle, PLSS provides broad geometric anchor points within which more complex parcel systems can exist.

PLSS is really only used in defining federal ownership boundaries, and it is generally not used inside of more developed areas, where assessor parcels are the dominate framework for defining land ownership – its use tends to mainly focus on rural and remote areas.

Ideally, the PLSS and county assessor boundaries would agree completely, but in practice that has not often been the case in California. This leaves a complicated choice for how to apply PLSS in CPAD. The general approach that GreenInfo has taken has been to not align any data to PLSS – if federal BLM or Forest Service data is itself PLSS-aligned, then we have focused on deciding whether to align those ownerships to parcels. Where PLSS, the owning agency and parcels disagree, there is little cause to choose PLSS over the other two sources. There are also cases where federal agency lines and parcels agree, and PLSS is different.

AERIAL IMAGERY – A CPAD Reference

Aerial imagery is a valuable reference for developed CPAD holding geometry where no GIS source information is available, for reviewing holdings for accuracy (parcels or ownership boundaries are not always correct in relation to what is actually on the ground), and for interpreting various CPAD attribute determinations (e.g., public access

and hardscape areas in smaller urban park locations). GreenInfo uses current year National Agricultural Imagery Program (NAIP) and online Bing (through Esri services) imagery for this purpose. Both sets of imagery are considered reasonably accurate for CPAD purposes.

It is worth noting, however that sometimes even aerial images that appear aligned with streets and other data may suffer from overall error, though CPAD editing efforts are generally not able to distinguish this. And, in some of these cases, it appears that there are occasional rural roads that were built on incorrect ownership lines – looking at an air photo might suggest the boundary line (PLSS or assessor) is misplaced, but in fact those lines are correct and the road is actually in the wrong place!

As a general practice, GreenInfo now limits use of air image reviews to relatively limited circumstances, instead favoring assessor or source agency data lines.

RECOMMENDATIONS FOR ALIGNMENT BASE DATA

1. CPAD should continue to be developed using assessor boundaries as its primary base alignment layer – much of CPAD in metropolitan areas is parcel aligned now as are large swaths of rural areas as well (but not all counties). This guideline is consistent with the move of California’s state and local agencies (plus the private sector) to increasingly use GIS-based parcel data in a wide range of business practices.

2. CPAD attributes should continue to include information on what data was used for each holding’s boundary alignment (improvements to the CPAD database provided better tools for doing this as of early 2013).

3. Parcel data used in CPAD should be as current as possible, given constraints on availability and resources to acquire it – in general, annual updates of each county’s parcel data should be undertaken. The State should make every effort to provide whoever manages CPAD with access to any current, statewide parcel data.

4. The use of assessor base data should be conditioned by the following:

a. Where owning agencies/organizations confirm that their non-parcel aligned GIS data is based on accurate survey information and this information is not significantly contradicted by aerial imagery or other corollary data (including the general accuracy of the surrounding assessor parcel pattern), CPAD will seek to incorporate these boundaries on a case by case basis, as resources allow.

NOTE: This process is likely to be an ongoing one and need further discussion with select federal and state agencies.

b. Where source, assessor or PLSS lines diverge greatly from what shows clearly as a boundary line on authoritative aerial imagery, CPAD may be adjusted to fit the aerial image. This is primarily for urban or suburban areas, or for rural areas with highly distinct features (fence lines, etc.), where there is significant variance between base line data and the aerial image.

c. The CPAD managing organization should support others’ efforts to improve county assessor boundary accuracy and encourage resolution of PLSS and county assessor boundary differences.

d. In areas of broad federal land ownership that are generally aligned to PLSS boundaries (in particular, the southern California deserts), CPAD should strive to rely on the source agency, PLSS-aligned data.

II. DATA THAT FRAMES CPAD GEOMETRY

CPAD data must, to the greatest extent feasible, relate well to framing data – those layers of geographic data that are frequently used in relation to CPAD data. The following are the main types of this data:

- **County boundaries:** Boundary lines delineating the border of a particular county government's jurisdiction
- **Coastline(s):** Data defining where the accepted border is between coastal water and land.
- **Other hydrologic data :** Data that is used to define lakes and other interior water features of CPAD holdings
- **Roads:** Road data is used as an additional check on CPAD boundaries, particularly where it is created as polygon rights of way (rather than just street centerlines). Such data is usually part of an overall county parcel layer.

DISCUSSION

County Boundaries

County boundaries are used in CPAD to define Units (holdings under common ownership that lie within a county) as well as to interpret the geometry of some holdings where parcel or other core data is not available or is not accurate. GreenInfo uses the official state county file, titled **cnty2409_1**, which is managed by the Calif. Department of Forestry and Fire Protection (Cal Fire) and described in its metadata as:

In late 1996, the Dept of Conservation (DOC) surveyed state and federal agencies about the county boundary coverage they used. As a result, DOC adopted the 1:24,000 (24K) scale U.S. Bureau of Reclamation (USBR) dataset (USGS source) for their Farmland Mapping and Monitoring Program (FMMP) but with several modifications. Detailed documentation of these changes is provided by FMMP and included in the lineage section of the metadata. A dataset named cnty24k97_1 was made available (approximately 2004) through the California Department of Forestry and Fire Protection - Fire and Resource Assessment Program (CDF - FRAP) and the California Spatial Information Library (CaSIL).

In late 2006, the Department of Fish and Game (DFG) reviewed cnty24k97_1. Comparisons were made to a high-quality 100K dataset (co100a/county100k from the former Teale Data Center GIS Solutions Group) and legal boundary descriptions from (<http://www.leginfo.ca.gov>). The cnty24k97_1 dataset was missing Anacapa and Santa Barbara islands. DFG added the missing islands using previously-digitized coastline data (coastn27 of State Lands Commission origin), corrected a few county boundaries, built region topology, added additional attributes, and renamed the dataset to county24k.

In 2007, the California Mapping Coordinating Committee (CMCC) requested that the California Department of Forestry and Fire Protection (CAL FIRE) resume stewardship of the statewide county

boundaries data. CAL FIRE adopted the changes made by DFG and collected additional suggestions for the county data from DFG, DOC, and local government agencies. CAL FIRE incorporated these suggestions into the latest revision, which has been renamed cnty24k09_1.

The official county boundary file is clipped to the coastal shoreline, but an additional file (from Esri StreetMap) is available with full-extent county boundaries (i.e., boundaries that extend to full legal jurisdiction in the ocean areas). This file also contains a variety of feature types including:

- Full state outline, not interior county lines
- State with county boundaries
- State with county multi-part boundaries
- State and counties as polylines

Coastline/Shoreline

Coast and shore line data is useful in defining the land extent of holdings in CPAD and in adjusting holding boundaries that do not have authoritative data (e.g., are created from reviewing printed maps, etc.).

The official California coastline data set is from 1993 and is only generally accurate. GreenInfo has used a custom file based on county boundaries trimmed by shorelines from Esri but has edited this file extensively through use of aerial imagery so that it is now a highly accurate depiction of shoreline as defined visually, particularly in the Bay Area (areas around Long Beach Harbor need some additional editing for tidal waterways).

Efforts by NOAA and the Coastal Conservancy are underway to improve this depiction, but GreenInfo has also begun developing a very accurate coastline that uses extremely precise geometry from Open Street Map. As we progress with CPAD we wish to bring that data into the state county boundary file. This coastline data is mostly used for cartography, but is sometimes used for analysis. While experts may work through very fine grained details such as exact locations of high/medium water marks, what is most needed is a generally accurate (to air photos) alignment framework.

Other Hydrologic Data

Inland waters (lakes, reservoirs, etc.), as well as bays and oceans are coded as "Water" for CPAD holdings when they are 10 acres or greater. This is done using a slightly modified version of the National Hydrology Dataset (NHD). CPAD delineates some "Water" holdings that are smaller than 10 acres where they were previously recorded or where resources allow for better precision.

Roads

Road line data is useful as a general reference in CPAD work, but given that many parcel data sets define road rights of way (polygons) and that aerial imagery often matches these rights of way, our current centerline based files are of limited use.

RECOMMENDATIONS FOR REFERENCE DATA

1. Improve the California County boundary file by integrating it with GreenInfo's shoreline geometry. Provide this data set to CAL FIRE as recommended next generation county file. Consider also recommending the addition of a file with full extents of county jurisdictions out to the ocean and bays.
2. Continue using current practices for inland lakes and other water bodies, but explore options and directions about this approach in the work on CPAD database design.

III. DATA STANDARDS USED IN RELATION TO BASE DATA

In addition to base geographic data, there is the matter of standards that affect overall data alignment. The most important of these is how data is projected.

DISCUSSION

Based on long-standing state approaches, CPAD uses the California Teale Albers, NAD83 projection:

```
+proj=aea  
+lat_1=34.00  
+lat_2=40.50  
+lat_0=0.00  
+lon_0=-120.00  
+x_0=0.000  
+y_0=-4000000.000  
+ellps=GRS80  
+units=m  
+datum=NAD83
```

Source data for CPAD, however, comes in many different projections, including the seven California State Plane Zones, UTM zones 10 and 11, and national projections for federal data. Care is taken to use the coordinate system transformation that is appropriate in each case.

RECOMMENDATIONS

1. We recommend that CPAD continues to be maintained in California Teale Albers, NAD83, particularly since it provides the least distortion of area calculations across the state.