Appendix K
Calibration Targets and Criteria

1.0 Purpose

The purpose of this appendix is to present the development of the calibration targets and criteria developed as part of the South Platte Decision Support System (SPDSS) Alluvial Groundwater Model. In the SPDSS there was a specific task, Task 48.2 Development of Calibration Targets and Criteria to define the calibration process, define the calibration periods (steady-state, transient and validation), and to develop calibration targets and criteria to be used for the SPDSS Alluvial Groundwater Model. The objective of this task is as follows:

> *To define the calibration process, calibration periods and develop a set of calibration targets and criteria to be used to assess the calibration of the alluvial groundwater model being developed under Task 48.*

In this technical memorandum, the model calibration process was defined for the SPDSS Alluvial Groundwater Model. The parameters used for assessing model calibration and model parameters to be modified during model calibration were defined. A set of calibration targets and criteria both numeric and non-numeric were defined in order to assess the calibration of the SPDSS Alluvial Groundwater Model.

A copy of this technical memorandum is included in this Appendix K.

Task 48.2 Development of Calibration Targets and Criteria

SPDSS Phase 4 Task 48.2 Development of Calibration Targets and Criteria - Final

To:	Ray Alvarado, CWCB
From:	Camp Dresser and McKee Inc. Mike Smith, Mark McCluskey, Bill Fernandez, and Gordon McCurry
Subject:	SPDSS Groundwater Component Phase 4 Final Task 48.2 Technical Memorandum, Development of Model Calibration Targets and Criteria, South Platte Alluvial Groundwater Model
Date:	October 8, 2008

Introduction

Phase 4 Task 48 of the South Platte Decision Support System (SPDSS) includes calibrating a groundwater flow model of the alluvial aquifer system within the South Platte Alluvium Region within Water Division 1. The model includes unconsolidated alluvial deposits of the South Platte River mainstem, extending downstream from Chatfield Reservoir to the Nebraska state line at Julesburg. In addition, the model includes unconsolidated alluvial deposits of the major tributaries to the South Platte River downstream of Chatfield Reservoir.

This Technical Memorandum (TM) was undertaken under Task 48.2 of Phase 4 of the SPDSS, to develop calibration criteria, including selection of field data (targets) to be used during the model calibration. This TM summarizes the methodology and data that are anticipated to be used in the model calibration process.

Approach

Calibration targets and calibration criteria have been developed for both the steady-state and transient model simulations for the SPDSS alluvial groundwater model. A general description of model calibration and the process that will be used to calibrate the SPDSS model are described in Sections 1 and 2, respectively, of this TM. Several types of data are used as targets in the model calibration; using multiple targets increases the confidence that the model accurately represents the stresses imposed on it. The calibration targets and the periods used to represent the steady-state and transient calibration periods are discussed in Section 3. The numeric values (criteria) that will be used to evaluate how well the model is calibrated are described in Section 4. Sections 5 and 6 provide a summary and recommendations, respectively. The following table summarizes the sections contained in this TM.

Section	Description
1.0	Model Calibration Overview
2.0	Calibration Process
3.0	Selection of Calibration Targets
3.1	Calibration Time Periods
3.2	Groundwater Level Targets
3.3	Streamflow Targets
3.4	Stream Gain/Loss Targets
3.5	Other Targets
3.6	Development of Target Database
4.0	Selection of Calibration Criteria
4.1	Target Criteria
4.2	Objective Function
4.3	Weighting Factors
5.0	Summary and Conclusions
6.0	Recommendations

1.0 Model Calibration Overview

Model calibration is an iterative process where selected sensitive model parameters are adjusted within predetermined ranges until the simulation results match observed data to an acceptable degree (Anderson & Woessner 1992). A calibrated model simulates historical conditions within an acceptable range of uncertainty, thus representing the effects of past inflows and outflows. Calibration is by its nature non-unique, i.e. many combinations of model parameters may results in a model that fits the field data. If sufficient data exist, calibrated models may be used to simulate other historical time periods in order to verify that the model is able to adequately represent the system response for a period that was not used for the calibration. This process is termed model validation. Calibrated models that adequately simulate the validation period are considered to be even more robust in their ability to reproduce the processes that have occurred historically in the area modeled. Models are most commonly used for evaluating future conditions. A calibrated model may be used to estimate future and past conditions with confidence if the stresses imposed on the model are comparable to those imposed during the calibration and validation period.

The field data used for comparison against simulation results are termed calibration targets in this TM. Calibration targets are defined in terms of the type of measurement, its location and date of measurement, and measurement value. The SPDSS alluvial groundwater flow model will be calibrated using the following:

- Water levels observed at wells
- Stream Flow, Diversions and Gain-Loss
- Groundwater Budget, including groundwater evapotranspiration (ET)
- Other Non-Numeric Data (e.g. dry cells, flooded cells)

The calibration criteria define the acceptable differences between the measured and simulated values for each calibration target. The relative importance of the calibration targets will be incorporated through weighting factors assigned to each target class. The calibration targets and their weighting factors are discussed in Section 3.

The calibration process used in the SPDSS alluvial groundwater model involves the use of both manual and automated parameter estimation techniques. The automated parameter estimation technique is an alternative to manual methods in which a parameter is modified, a simulation made, the results are evaluated, another parameter is modified and the process repeats. The size and complexity of the SPDSS model warrants an automated process, once a stable model producing a water budget within a reasonable range is obtained and initial sensitivity analyses are completed. Automated parameter estimation techniques have been employed for over a decade and extensive guidance exists for their implementation (Doherty 2004). The automated model calibration will be undertaken as described in Section 4.

2.0 Calibration Process

Model calibration is the process in which model input parameters are varied within predetermined ranges in an iterative manner until modeled results match observed data within an acceptable range. The calibration process is a series of steps undertaken to calibrate a model. This process is illustrated in Figure 2-1.

The overall model calibration process will be conducted in three steps.

1. Calibration to a representative steady-state period
2. Calibration to a representative transient period
3. Verification of calibration to the full study period

The time periods for these calibration steps are described in Section 3.1 below. The calibration process starts with the development of model input files. This includes defining the model configuration, initial model parameters, and stresses. These initial inputs are developed using the data centered process using programs and tools developed specifically for this purpose. These tools are discussed further in the Task 48 Alluvial Groundwater Modeling Report. This initial model must be assessed to ensure that it is numerically stable over likely ranges of input parameters, and that factors such as dry cells are minimized. Model control parameters, including selection of a solver and appropriate solution parameters must be configured to enhance this stability. Initial sensitivity analyses are conducted in order to identify model parameters that have the greatest control over the goodness of fit of the model to field data.

The calibration process will proceed by first approximating model parameters using a steady-state calibration period (Figure 2-1). The model parameters from the steady-state calibration will then be used as initial estimates for the transient calibration period to refine the model. Finally, the calibrated model will be run over the entire study period to verify that acceptable agreement between the model and field data has been reached. Each of these steps is iterative in nature. This is the process that will be implemented for the SPDSS and is described in more detail below.

Automated parameter estimation techniques will be used during each stage of the calibration process. Since the goodness of fit of the model is defined by comparing model results to field data, a quantitative measure of this fit needs to be developed. This measure is defined as an objective function. An objective function is used by parameter estimation programs such as MODFLOW-2000 and PEST to quantify the model's goodness of fit when using either manual techniques or a parameter estimation program. The objective function that will be used for the SPDSS model will include a weighted combination of heads and fluxes, as discussed in Section 4.

The model parameters being modified during calibration are those that, based on the available data, have the largest uncertainty and impact the objective function value as they are varied. They include the following, listed in the likely order of priority for implementation and decreasing order of parameter uncertainty:

- Aquifer hydraulic conductivity (K)
- Streambed conductance
- Recharge
- Well pumping
- Lateral boundary inflows
- Specific yield

Numerous simulations using the parameter estimation program will be run by varying these model parameters within predefined ranges that bracket the probable range in these characteristics. The MODFLOW model results will be evaluated using the automated parameter estimation package according to the calibration criteria discussed in Sections 3.1 and 3.2. These criteria will be evaluated using an objective function that quantifies the difference between model simulated and field observed values (residuals). For each simulation, a residual of the weighted calibration targets will be produced using the objective function. Each simulation will also be evaluated by its ability to meet the target water budget, head convergence and mass balance criteria. Typically, head convergence will be acceptable using a value of 0.1 ft for transient simulations, and a mass balance of less than one percent for final time steps in a stress period. There are no generally accepted standards for head convergence criteria or mass balance, however, the selected criteria are adequately conservative and will result in a valid numerical solution.

The calibration will proceed using a combination of automated and manual methods to obtain the best fit between model results and field data. The automated calibration process will provide information on the residuals for the parameters used in the objective function. These parameters include:

- groundwater levels,
- stream flow,
- stream gain/loss.

However, the automated calibration process can lead to several combinations of model parameters resulting in a similar residual. To further evaluate calibration results other parameters will be evaluated manually. These parameters are both numeric and non-numeric and include the following:

- model water balance,
- groundwater budget,
- stream diversions,

- the location of groundwater ET,
- the number and distribution of flooded and dry model cells,
- the spatial distribution of residuals from the objective function.

Each of these targets are described in more detail in Section 3.

Steady-state and transient model simulations in a series that result in the best fit (i.e. lowest objective function value) will be evaluated against other non-numeric calibration criteria not included in the objective function. The set of input parameters that result in simulations that meet the calibration criteria will be used as inputs for the model validation time period, and the results will be evaluated. If the criteria are not met for the validation period (January 1950 to October 2006) then the model inputs and model objective function will be re-evaluated and the model simulation process will begin again (Figure 2-2). The model will be considered calibrated when the calibration criteria are met over the validation period.

The last step in the modeling process is a formal parameter sensitivity analysis that describes the calibrated models sensitive to various parameters. During this formal sensitivity analysis, calibrated values for hydraulic conductivity, streambed hydraulic conductivity, recharge, and boundary conditions will be systematically varied within acceptable ranges. The results of a sensitivity analysis allow one to evaluate the effect of individual parameters on the calibrated model. This analysis will help identify the parameters that could benefit from additional data collection in the future, which would help reduce model uncertainty.

3.0 Selection of Calibration Targets

This section describes the general methodology and steps used to identify and select the field data that will be used for calibrating and validating the Phase 4 alluvial groundwater model.

The selected time periods for the steady-state, transient and validation periods are presented in Section 3.1. The groundwater level measurement (head) targets are defined and discussed in Section 3.2. Streamflow targets are discussed in Section 3.3. Stream gain/loss targets used for model calibration are discussed in Section 3.4. Targets that will be used in a manual calibration step, including water mass balance, groundwater budget, diversions and locations of flooded and dry cells, are discussed in Section 3.5. The methods used to develop the calibration target database are discussed in Section 3.6.

3.1 Calibration Time Periods

In order to increase the efficiency of the calibration process, time periods shorter than the full study period of 1950 to 2006 were selected for the initial calibration. For the steady-state calibration period (1991-1994), a period with relatively steady or consistent conditions was selected. For the transient calibration period (1999-2005), a period containing years with large variation in climatic and streamflow conditions was selected. This section describes how these calibration time periods were selected.

Streamflow, precipitation and water level data were used in this analysis. Pumping data were not readily available, so the pattern of alluvial well development based on decrees, was used.

The following process was used to select both the steady-state and transient calibration time periods. A summary of the data are presented later in this section:

1. Daily data for representative precipitation and streamflow gages were downloaded from HydroBase. The average annual streamflow and precipitation for each year of the study period were evaluated at monitoring locations distributed throughout the study area.
2. To help select the steady-state calibration period, water levels in alluvial wells with long-term records were evaluated to identify periods with relatively minor water level change. By definition there is no change in aquifer storage in a steady-state period so water levels should show no long-term increases or decreases.
3. The relative level of pumping in the South Platte alluvial aquifer was estimated by querying the water rights database in HydroBase and summarizing the decreed pumping rates of high capacity wells (greater than 50 gpm) for each year.
4. The total number of wells with water level measurements within the study area was summed for each year. This identified years with more data, which could better support the calibration process. Years were selected for further evaluation if they had at least 300 wells with at least one measurement.
5. The number of wells with four or more groundwater elevation measurements for each of the years selected in the previous step was determined. Four measurements per year were used as a threshold to assist in evaluating seasonal groundwater level changes in the transient calibration time periods.
6. The spatial variability of groundwater elevation data for each year was assessed using plots of well locations and the number of water level data points at each of the observation wells. A goal was to select time periods with a uniform spatial distribution of wells that include multiple measurements in a year.

The process used to select the steady-state and transient calibration periods represented a balance of the individual datasets described above, using engineering judgment to make the final selection of periods. These periods were reviewed and agreed to by the State prior to finalizing the selections.

For the steady-state calibration period, the five-year period from 1991 to 1994 was selected as representative of relatively consistent conditions for flows, precipitation, number of new wells, and water levels within the alluvial system. For the transient calibration period, the seven-year period from 1999 to 2005 was selected. This period includes a wide range in flows and climatic conditions, including the record drought in 2002, and significant changes in the number of wells. For the model verification period, the entire time period (1950 – 2006) will be simulated. The following describes the available data and the hydrologic characteristics of the selected calibration time periods.

3.1.1 Water levels

Groundwater level data is one of the key datasets upon which model calibration will be based. The ideal time period for model calibration is one for which a set of measurements exists that cover the model area uniformly and over all time periods that are simulated. During the selected steady-state and transient time periods there is a reasonably good distribution of water level data, both spatially and temporally. Figures 3-1 through 3-3 summarize the spatial distribution of available water level data that will be used in each of the model calibration periods. An important aspect of selecting the transient calibration period is that it contains a large number of the wells that include at least four measurements in a given year in order to adequately evaluate seasonal changes in water levels. Figure 3-4 shows the number of wells that have at least four measurements per year. Many of these are from SPDSS-installed wells that are equipped with data loggers. As shown in the Figure 3-4, periods with a relatively large number of wells that fit this criterion (defined here as at least 20 wells) include 1967-68, 1974-78, 1989-94, and 2000-2006. The last two of these periods generally include over 50 wells per year. Table 3-1 lists the number of wells and measurements available for each period.

Table 3-1 Summary of Groundwater Measurements for Model Calibration

	Steady-state Calibration Period (1991 - 1994)	Transient Calibration Period (1999 - 2005)	Validation Period (1950 - 2006)
Number of wells measured	391	605	4,811
Number of measurements	3,205	23,826	75,049
Number of wells with at least 4 measurements in a year	82	170	431

3.1.2 **Streamflow**

Annual streamflow at three gages located throughout the model area were used to evaluate hydrologic conditions:

- South Platte River near Kersey (6754000) gage,
- South Platte River at Henderson (6720500) gage,
- Cache la Poudre River near Greeley (6752500) gage.

The data from these gages were compared to each other over the period of record; relatively wet and dry cycles at the three gages were found to be in general agreement, as shown on Figure 3-5. The average flow at each gage for various periods of record including each calibration period are shown in Table 3-2. The difference in average flows for each period of record is compared to the full period of record for each gage and was used in the evaluation.

The results in Table 3-2 show that average streamflow was higher during the study period (1950-2005) than the period of record average flow for each gage by 8 to 18%. The steady-state period has lower streamflow than the period of record average, by about 15%. The streamflow during this period is lower than both the full period of record and the study period but other criteria (such as the number of observation wells) support the use of the selected calibration periods. The transient period has streamflow ranging from -8 to 5% of the average flow for each gage. This period shows significant variability in streamflow (Figure 3-5), with 1999 showing very high flows at the Kersey and Henderson gages, 2002 showing some of the lowest flows on record, and 2005 having average flows.

Table 3-2 Streamflow Classification of Calibration Periods

	South Platte at Kersey		South Platte at Henderson		Poudre near Greeley	
Period of Record (POR)	(1901 to 2006)		(1926 to 2006)		(1903 to 2006)	
	Average streamflow (af/yr)	% Difference from POR Average	Average streamflow (af/yr)	% Difference from POR Average	Average streamflow (af/yr)	% Difference from POR Average
Entire Period of Record	640,400	---	313,100	---	98,100	---
Study Period [1950- 2006]	719,900	12%	457,100	8%	130,500	18%
Steady-state Calibration Period [1991-1994]	553,300	-14%	261,880	-16%	82,800	-16%
Transient Calibration Period [1999-2005]	586,600	-8%	327,600	5%	91,100	-7%

3.1.3 Precipitation

Annual precipitation data totals at three precipitation gages with long-term records that were spatially distributed in the study area were obtained and evaluated. These gages are:

- Byers 5 ENE,
- Fort Collins, and
- Denver Stapleton International Airport.

For each of the gages the period of record average was compared with proposed steady-state and transient calibration time periods. As shown in Table 3-3, there is relatively little variation in average precipitation between the period of record average and the calibration periods, with the steady-state period showing slightly higher precipitation and the transient period showing slightly less. The transient period includes some of the highest 20.68 in (Ft Collins, 1999) and lowest 7.68 in (Byers, 2002) annual precipitation amounts during the study period, as shown in Figure 3-6.

Table 3-3 Climate Classification of Calibration Periods

	BYERS 5 ENE.		FORT COLLINS		DENVER STAPLETON	
Period of Record	(1948 to 2006)		(1900 to 2006)		(1948 to 2006)	
	Average precipitation (in/yr)	Difference from POR Average	Average precipitation (in/yr)	Difference from POR Average	Average precipitation (in/yr)	Difference from POR Average
Entire Period of Record	14.96	---	15.13	---	15.47	---
Study Period [1950 to 2006]	14.96	0%	15.19	0%	15.45	0%
Steady-state Calibration Period [1991 to 1994]	15.29	2%	16.40	8%	15.45	0%
Transient Calibration Period [1999 to 2005]	14.71	-2%	15.15	0%	14.97	-3%

3.1.4 Aquifer Storage

Water level data from HydroBase were used to evaluate alluvial aquifer storage. Water level data from alluvial wells located throughout the active model domain were examined, using information developed under the Task 39 and Task 44 evaluations. The hydrographs for a subset of these wells is shown in Figure 3-7 to illustrate the long term changes in alluvial groundwater levels. As concluded in the Task 44.3 TM and indicated in Figure 3-7, alluvial aquifer water levels show little change from year to year in most locations. Although some wells show declines and others show increases in water levels, the majority of alluvial wells show relatively stable levels compared to the seasonal changes. This indicates that, on an annual basis, there is little change in alluvial aquifer storage and so almost any time period could be used for the steady-state period based on this criterion. However, most of the hydrographs reviewed including many shown in Figure 3-5 and others included in the Task 39 and Task 44.3 TMs show significant seasonal fluctuations. These water level changes are a focus of the transient calibration efforts.

3.1.5 Well Development

Alluvial well development in the South Platte was evaluated using the annual total for decreed high capacity wells (those with decreed pumping rates greater than 50 gpm). Periods of little change in decreed capacity are preferred for the steady-state calibration period while periods of large change in decreed capacity would be preferred for the transient period.

Using a water rights query from HydroBase, the decreed capacity and adjudication date of wells in the South Platte alluvium within water districts that are included at least partially within the study area were identified for each year (Figure 3-8). Well decreed capacity in the South Platte started to increase in the mid-1930's and continued at a high rate of annual change through the 1960's. After 1970 the capacity of newly decreed high capacity wells decreased to annual numbers comparable to before the mid-1930's except for isolated years in 1985, 1990, 2002 and 2003 (Figure 3-8). This suggests that pumping stresses have been relatively uniform in the alluvial aquifer since the early 1970's with the exception of the four years noted.

3.1.6 Conclusion

The data reviewed and discussed in the previous portions of Section 3.1 supports the selection of average values from 1991 through 1994 for the steady-state calibration period, and monthly values from 1999 through 2005 for the transient calibration period.

The following sections summarize the data that will be used as the steady-state and transient calibration targets.

3.2 Groundwater Level Targets

There will be two categories of groundwater level targets based on the accuracy of their measurement. Water levels in wells are measured as a depth to water below some measuring point, which is the top of the well. The water level measurement is converted to a groundwater level elevation by subtracting the depth to water from the measuring point elevation.

The first category of groundwater level targets is data from wells that have had their measuring point elevations surveyed. Data from these wells have accurate measuring point elevations which provide accurate groundwater elevation readings. These data will be included in the calibration process through their water level elevations.

Groundwater level data from most wells measured within the South Platte River alluvial aquifer fall into the second category of targets. These wells have measuring point elevations in HydroBase that are estimated from topographic maps. The use of topographic maps to estimate elevations results in significant uncertainty in the elevation of the water surface in the well. This uncertainty is equal to plus or minus half of the contour interval for the map used for the estimate. Typically the map contour interval was 20 feet resulting in an uncertainty of plus or minus 10 feet for most of the study area. However, the accuracy of the depth to groundwater measurements is very high, typically less than 0.1 feet error. To take advantage of this measurement accuracy and incorporate these data into the calibration process, the change in depth to water rather than the water level elevations will be used for wells that have not been surveyed. Depth to water data from surveyed wells will also be added to this dataset to provide a larger dataset of the same types of measurement data. This data set will reflect seasonal and year to year variation in water levels that will be used in transient calibration. Weighting of the water level data is discussed in Section 4.3.

The locations and numbers of wells for which groundwater level target data exist are shown in Figures 3-1 through 3-3 and Table 3-1, respectively, using the 06-01-2008 version of HydroBase. Model results will be compared against all measurements collected during the calibration and validation periods.

3.3 Streamflow Targets

Recorded flows at all major stream gages within the model domain were selected as targets due to their good spatial distribution and relatively complete periods of record within the study period. These gages were also used in the Task 46 stream gain/loss study so there is a consistent set of flow and gain/loss information for these gages and reaches. Simulated streamflow will be compared to observed data from 9 gaging locations shown on Figure 3-9.

These gages represent flow conditions over the mainstem of the South Platte River and also at two major tributaries within the active model domain. Details on the gage locations, periods of record, filling of missing records and other characteristics of these stream gages can be found in the Task 2 TM (LRE 2006). A summary of the streamflow values used as targets are provided in Table 3-4 below. The transient calibration and validation periods have a monthly set of individual flow values that will serve as the targets for these calibration periods.

Table 3-4 Average Streamflow (af/yr)[1]

Streamflow Monitoring Site and Number	Steady-state Calibration Period (1991 - 1994)	Transient Calibration Period (1999 - 2005)	Validation Period (1950 - 2006)
Plum Creek near Louviers[2] (6709500)	13,040	22,460	23,910
South Platte River below Chatfield[2] (6708000)	49,630	73,900	120,260
South Platte River at Denver (6714000)	158,660	199,230	232,550
South Platte River at Henderson (6720500)	262,260	327,460	336,880
South Platte River at Fort Lupton (6721000)	286,890	334,700	354,260
South Platte River near Kersey (6754000)	553,490	586,810	720,120
South Platte River near Weldona (6758500)	358,610	395,560	504,230
South Platte River at Balzac[4] (6759910)	289,060	307,170	415,840
South Platte River at Julesburg[3] (6764000)	302,830	257,910	418,740
Cache la Poudre River at Canyon Mouth near Fort Collins[2] (6752000)	191,980	183,290	226,030
Cache la Poudre River near Greeley (6752500)	82,590	91,280	115,190
Cherry Creek near Franktown[2] (6712000)	4,350	6,520	6,520
Cherry Creek at Denver (6713500)	16,660	27,530	17,390

[1] Values in acre-feet per year, rounded to nearest 10 units

[2] Represents stream flow into the alluvial model domain

[3] Represents stream flow out of the alluvial model domain

[4] The South Plate near Balzac gage has been combined with the South Platte at Cooper Bridge due to the gage being moved during the full simulation period

3.4 Stream Gain/Loss Targets

As discussed in Section 2.0, streamflow and stream gain/loss will serve as a calibration target during the automated parameter estimation process. A groundwater flow model is considered much more reliable when it is calibrated to more than one type of target, such as flux (in this case streamflow and stream gain/loss), in addition to head (Anderson & Woessner 1992). Observed streamflow (from gage records) and stream gain/loss reflect monthly diversions, tributary inflows, municipal discharges, and surface water gain/loss processes in the reach upstream of the gage. These flows can change significantly from month to month in the South Platte River basin and therefore are an important target to include in the model calibration.

Average gain/loss targets for each study reach are shown in Table 3-5 below. These values are from the Task 46 Stream Gain/Loss Estimates TM (CDM 2008). The negative values indicate that flow is from groundwater to the receiving streams. As with the stream flow data presented in the previous section, the transient calibration and validation periods have a monthly set of individual gain/loss values that will serve as the targets for these calibration periods.

Table 3-5 – Average Stream Gain/Loss (af/yr)[1]

Stream Gain/Loss Reach	Steady-state Calibration Period (1991 - 1994)	Transient Calibration Period (1999 - 2005)	Validation Period (1950 - 2006)
South Platte River, Chatfield to Denver	-68,100	-66,650	-55,060
South Platte River, Denver to Henderson	-63,750	-68,100	-32,600
South Platte River, Henderson to Ft Lupton	-45,640	-30,430	-27,530
South Platte River, Ft Lupton to Kersey	-177,490	-155,760	-158,660
South Platte River, Kersey to Weldona	-125,330	-99,980	-30,430
South Platte River, Weldona to Balzac	-95,630	-77,520	-58,680
South Platte River, Balzac to Julesburg	-212,270	-171,700	-185,460
Cache la Poudre River, Canyon to Greeley	-124,610	-73,900	-104,320
Cherry Creek, Franktown to Denver	-13,040	-21,010	-11,590

[1] Values in acre-feet per year, rounded to nearest 10 units

3.5 Other Targets

The automated calibration process using the targets discussed in Sections 3.2 through 3.4 could potentially lead to several combinations of model parameters resulting in a similar residual in the objective function. A further evaluation of the calibration results will be undertaken manually using other targets that are discussed in this section. These targets include model water balance, groundwater budget, stream diversions, the location of groundwater ET, the number and distribution of flooded and dry model cells, and the spatial distribution of residuals from the objective function. These targets will be evaluated using a combination of numerical criteria and engineering judgment. The numerical criteria are listed in Section 4.1.

3.5.1 Model Water Balance

The model water balance is the model output of the difference between simulated inflows and outflows. In MODFLOW-2000 the water balance is presented for individual stress periods and for the cumulative simulation. It is an indicator of the validity of the numerical solution. Only simulations that exhibit a mass balance error of 1 percent or less will be used in the analysis.

3.5.2 Preliminary Groundwater Budget

A groundwater budget is the compilation of the estimated inflows and outflows for the groundwater system being evaluated. The budget terms are estimated from available data and engineering judgment for those components with insufficient data available. These inflows and outflows must balance, thus a term with the greatest uncertainty is used as the balancing quantity to obtain this equivalence between inflows and outflows. Other terms with significant uncertainty will potentially be adjusted during the model calibration process. In the calibrated model, a volumetric water balance will be obtained, since this is the basis for a simulation. If the model calculated flow for a component that has been selected as a calibration target deviates from the calibration target value, then model parameters will be modified to obtain a match within the selected criteria level.

Estimates of the individual water budget components have been made based on data collected from a variety of sources as part of the SPDSS, including flows presented in published TMs, unpublished calculations made by the Consumptive Use Contractor (Leonard Rice Engineers), and unpublished calculations made during Task 48. All calculated inflows and outflows used in the calibrated model will be documented in the Task 48.1 Stress Inputs TM. Preliminary values estimated for the steady state calibration period (1991-1994) and the sources of the estimates are shown in Table 3-6.

The groundwater budget terms with the most uncertainty include estimated discharge to streams, reservoir leakage, subirrigation and changes in storage. Estimated discharge to streams has a relatively large uncertainty because stream gage records do not exist which would allow this term to be calculated for a number of tributaries. Reservoir leakage has a relatively high level of uncertainty because leakage estimates have been calculated using generalized soil characteristics and reservoir storage estimates since insufficient data are available for direct estimates. Subirrigation has a relatively large uncertainty because no data exist and the value is based on an estimated percent of the maximum potential subirrigation (difference between crop water requirement and all water supplies) computed in StateCU. The change in groundwater storage, typically not included in a steady state water budget, has a relatively large uncertainty because it was estimated using the preliminary groundwater model by calculating the additional water needed to maintain water levels in these tributaries at stream elevations. Therefore this term includes both storage declines in some designated basins and alluvial tributaries as well as other uncertainties associated with an un-calibrated preliminary model. Accordingly, the water budget presented in Table 3-6 should be considered preliminary.

Table 3-6 Preliminary Groundwater Budget, Steady State Period (Average 1991 -1994)

Inflows		Flux (af/yr)[1]	Source[2]
	Precipitation - based Recharge	70,100	StatePP, using Jan '08 unpublished LRE StateCU output
	Irrigation/Canal Recharge [3]	1,475,700	StatePP, using Jan '08 unpublished LRE StateCU output
	Bedrock Aquifer Flux	19,400	USGS calibrated model (5/30/2008)
	Lateral Boundary Inflow	22,200	CDM calculations
	Alluvial Underflow	27,700	CDM calculations
	Reservoir Leakage	31,400	CDM calculations
	Change in Storage [4]	173,300	Estimated based on preliminary groundwater model simulations
	Total	***1,819,800***	
Outflow	Agricultural Pumping[9]	531,700	StatePP, using Jan '08 unpublished LRE StateCU output
	Municipal & Industrial Pumping	49,300	CDM calculations, revised from Task 41.3 TM
	Bedrock Aquifer Flux	5,500	USGS calibrated model (5/30/2008)
	Alluvial Underflow	13,800	CDM calculations
	Evapotranspiration [5]	163,200	Jan 2007 Memo from Groeneveld and Prescott,
	Subirrigation [6]	72,700	StatePP, using Jan '08 unpublished LRE StateCU output
	Calculated Discharge to Streams [7]	926,300	Task 46.2 TM, Table 9
	Estimated Discharge to Streams [8]	57,300	Closure term of the preliminary groundwater budget
	Total	***1,819,800***	
	Net	**0**	
Difference		**0.0 %**	

[1] Values in acre-feet per year, rounded to nearest 100 units
[2] All calculated groundwater budget terms will be updated and presented in the Task 48.1 Stress Inputs TM to be prepared in Phase 5
[3] Includes canal leakage, deep percolation of surface water and deep percolation of ground water over the active groundwater model domain (SPDSS StatePP Update, July 7, 2008)
[4] The 'Change in storage' term represents water level declines and uncertainties in some flux terms prior to obtaining a calibrated model
[5] Evapotranspiration is from groundwater in areas mapped as phreatophytes within the active model domain
[6] Value presented is 25 percent of the maximum potential value computed in StateCU
[7] Discharges to streams calculated only for the South Platte River, Cache la Poudre River and Cherry Creek
[8] Other streams within the active model domain may be sources of additional leakage/discharge, but have insufficient data available to estimate these terms; this is used as a closure term to balance the groundwater budget
[9] Includes irrigation pumping over the active groundwater model domain (SPDSS StatePP Update, July 7, 2008)

3.5.3 Other Non-Numeric Calibration Targets

Other information will be used to evaluate model calibration. These are termed non-numeric criteria for the purposes of this TM, since they will not be included as part of the model calibration objective function. Each of the following targets will be evaluated using engineering judgment to assist in identifying the best combination of model input parameters. In some cases the targets will be evaluated against quantitative metrics, discussed in Section 4.1.

- Random geographic distribution of head, streamflow and stream gain/loss residuals
- Comparison of simulated and observed diversions
- Number and location of flooded and dry cells
- Correspondence between simulated and mapped areas of groundwater evapotranspiration

These types of data are not readily suited for automated parameter estimation techniques. Some of the targets listed above are estimated or derived values that have a higher level of uncertainty than the measured parameters of water level (head) and streamflow. In addition, most of the targets listed above are affected by changes in head and streamflow, so including them in the automated parameter estimation process would be redundant. These non-numeric targets will provide another means to evaluate how well calibrated is a given model run. They are considered a lower priority than the targets used in the automated parameter estimation process described in previous sections but still important elements of the model calibration.

3.6 Development of Target Database

Water level and stream flow data used in the calibration and validation processes resides in HydroBase and will be extracted and processed into a format suitable for use in the calibration process. The stream gain/loss information will be compiled from the Task 46 TM. The data used for calibration will include:

- Water elevation observations at specified well locations,
- Streamflow observations at specified streamflow gaging stations,
- Stream gain/loss.

4.0 Calibration Criteria

The objective of the calibration process is to obtain acceptable agreement between model calculated values and their corresponding measured values. The calibration process systematically varies model parameters within predetermined ranges based on site data and engineering judgment to obtain this agreement. Typically, the model parameters with the greatest uncertainty, including those that are not easily measured or can have significant spatial variability (such as aquifer hydraulic conductivity) are used for adjustment in calibration. The input parameters that will be adjusted during the calibration process are discussed in Section 2.

Calibration criteria are defined in this TM as the threshold values used to determine whether the model is sufficiently accurate. The criteria are the maximum differences between simulated and observed values, called the residuals. These are defined prior to beginning the calibration process and are considered as goals based on the modeling objectives and measurement or other errors inherent in the field measurements. Calibration criteria are defined for each type of calibration target. The criteria for each calibration target are discussed in Section 4.1. In addition to the evaluation of residuals, the model calibration process will also include maps, histograms, and scatterplots of residuals to help identify spatial trends.

The head and flow criteria are compared to the results provided in the automated model calibration package (PEST; Doherty, 2004) through its parameter estimation objective function, discussed in Section 4.2. The head and flow data used for calibration have varying degrees of uncertainty. This uncertainty in the observed values is factored into the calibration process by assigning weighting factors. Weighting factors that will be assigned to each calibration criterion are discussed in Section 4.3.

4.1 Target Criteria

Calibration targets have been developed for use with the automated parameter estimation objective function ('numeric' targets, described in Sections 3.2 through 3.4) and for the manual evaluation of model calibration using other targets ('non-numeric' targets, described in Section 3.5). The model calibration will first rely on the numeric targets and their criteria, with simulations meeting these criteria in the objective function being further evaluated with the non-numeric targets and their criteria to determine if additional calibration is needed. Each set of targets and criteria are presented in the sections below.

4.1.1 Criteria for Head and Flow Targets

The calibration criteria are listed below. They are considered to be reasonable calibration goals based on review of the data and engineering judgment but may need to be revised based on the initial model calibration efforts. There are no generally accepted standards for an acceptable agreement between modeled and observed heads and flows in guidance documents from organizations such as the ASTM or various state agencies.

- The absolute value of the mean difference of head residuals for wells with surveyed elevations across all time periods will be less than 5 feet.
- The absolute value of the difference between observed heads at surveyed wells across all time periods will be less than 5 feet for 75 percent of the observations.
- Observed and calculated average annual streamflows will be within 10 percent for all flows greater than 25,000 acre-feet/year.
- Observed and calculated average annual stream gain/loss within a given reach will be within 10 percent for all flows greater than 25,000 acre-feet/year.

4.1.2 Criteria for Other Non-numeric Targets

- Residuals for head and flows should be randomly distributed on a geographic and temporal basis.
- The model volumetric water balance calculated by the model will be within 1 percent for the steady-state calibration simulation and within 5 percent at the end of the transient simulations.
- The residual for each of the components of the groundwater budget will be less than ten percent for 75 percent of the budget components.

- The absolute difference between simulated and observed diversions across all time periods will be less than 10 percent of observed diversions.
- The number of flooded and dry cells will be less than one percent of the model cells in the active model domain and will be randomly distributed.
- The difference between simulated and mapped areas of groundwater evapotranspiration will be less than 10 percent when examined at the end of the transient simulations.

4.2 Objective Function

The automated calibration process will utilize a combination of optimization techniques that minimizes the residual between model-calculated values and observed values for a corresponding set of model inputs (Doherty, 2004). Minimizing the residual involves use of an objective function that is comprised of a defined set of error terms for each target along with a weighting function for that target. There are three key parameters that constitute the calibrations targets: head, streamflow and stream gain/loss. The head is broken down further into three categories (surveyed wells, non-surveyed wells with multiple measurements, and non-surveyed wells with a single measurement) to allow different weighting factors to be applied to each, as discussed in Section 4.3. This objective function takes the following form:

$$OF = w_1(H_1) + w_2(H_2) + w_2(H_3) + w_3(S) + w_4(G)$$

Where:

OF is the Objective Function
H_1 is the sum of head residuals of water level elevations at surveyed wells
H_2 is the sum of head residuals of water level fluctuations at non-surveyed wells that have more than one measurement
H_3 is the sum of head residuals of depth to water at non-surveyed wells that have one measurement
S is the sum of streamflow residuals at stream gaging stations
G is the sum of stream gain-loss residuals at reaches between gaging stations
w_1, w_2, w_3, w_4, and w_5 are the weighting factors for each target

Each target may be subdivided into groups based on geographic location, data accuracy, location along the mainstem or a tributary, or other factors. In these cases there will be a separate target term and weighting factor in the objective function. The various measures must also be scaled so that the different error terms are comparable. For example, streamflows are in flow rates (cfs) that can vary by hundreds to thousands of units compared to water level elevations that are in length units (feet) and typically vary by tenths of units. The various terms in the objective function will be scaled to remove these unit dependencies. Further details on objective functions are provide in the PEST manual (Doherty 2004) and in Anderson and Woessner (1992), among other references.

4.3 Weighting Factors

Weighting factors will be utilized to develop the goodness of fit statistics for the model calibration by considering the reliability of individual measurements. The weighting factors represent an estimate of the measurement error. These errors are an estimate of the underlying accuracy of the measurement, not a measure of variation in the data over time. For example, depth to water measurements are typically estimated to be accurate within plus or minus 0.1 feet, while streamflows are estimated to be accurate within plus or minus 10 percent. Weighting factors also will be applied to account for factors such as clustering of observations in time or space in addition to applying scaling factors and the measurement error-based weighting factors.

A weighting factor will allocate the contribution of flow and head target types to the overall objective function. Currently, the relative weighting for the different classes of target data are as follows.

- Well measurements with survey elevations: 25 percent of total
- Wells that are not surveyed and more than 1 measurements in a calibration period: 20 percent of total
- Wells that are not surveyed and one measurement in a calibration period: 5 percent of total
- Stream gage annual flows: 15 percent of total
- Stream gain/loss: 35 percent of total

5.0 Summary and Conclusions

CDM has completed SPDSS Task 48.2, the development of calibration targets and criteria for the South Platte alluvial groundwater model. This TM describes the process used to identify the target locations and time periods for which observed data will be compared to model simulation results. This TM also discusses and presents the calibration criteria and the range in acceptable differences between observed and simulated measurements.

Below are conclusions from completion of this task:

- Model calibration will include steady-state and transient time periods.
    - The steady-state period will simulate average annual conditions for 1991-1994, which represents a period with relatively steady or consistent conditions. The results will be used as initial input to the transient (time-varying) calibration followed by an initial transient period of 1995-1998 conditions.
    - The transient period will simulate monthly conditions during the 1999-2005 period.
    - Following the calibration to these periods, a validation run will be completed that simulates the entire study period of 1950 through 2006.
- A combination of automated parameter estimation and manual techniques will be used to assist in calibrating the model. This is considered to be the most effective and efficient method to calibrate a model as large and complex as the South Platte River basin alluvial groundwater model and follows current modeling standards. Accordingly, an objective function will be used to determine goodness of fit for the parameter estimation simulations.
- Using automated techniques, the model will be calibrated on groundwater level elevation data, changes in water levels fluctuations over time, streamflow, and stream gain/loss. Additional criteria will be used to assist in evaluating model calibration including distribution of residuals, the ability to match the target ground water budget,

model water balance errors, presence of wet and dry cells, comparison of historic diversions, and locations of groundwater evapotranspiration.

- Parameter weighting will be used to address measurement errors in parameters used in calibration and to allow groundwater level and streamflow data to be evaluated on a similar numeric basis even though they have different measurement units and magnitudes.
- Model calibration will be undertaken by varying aquifer hydraulic conductivity, streambed conductance, recharge, well pumping, lateral boundary inflows and aquifer specific yield.
- The calibration process is expected to result in a model that can be used with confidence to simulate historic and future groundwater inflows and outflows that are occurring within the alluvial aquifer of the South Platte River basin.

6.0 Recommendations

The following are recommendations from the development of calibration targets and criteria:

- Continue current water level measurement programs, especially at wells that have data loggers and have been surveyed.
- Use the targets and criteria presented in this TM for calibrating the alluvial groundwater flow model in Phase 5.

References

Anderson, M.P. and Woessner, W.W. 1992. Applied Groundwater Modeling. Academic Press, Inc. San Diego. 381 pp.

CDM, 2008. Task 46 Stream Gain/Loss Estimates Technical Memorandum, prepared for Colorado Water Conservation Board and Colorado Division of Water Resources, April 10, 2008.

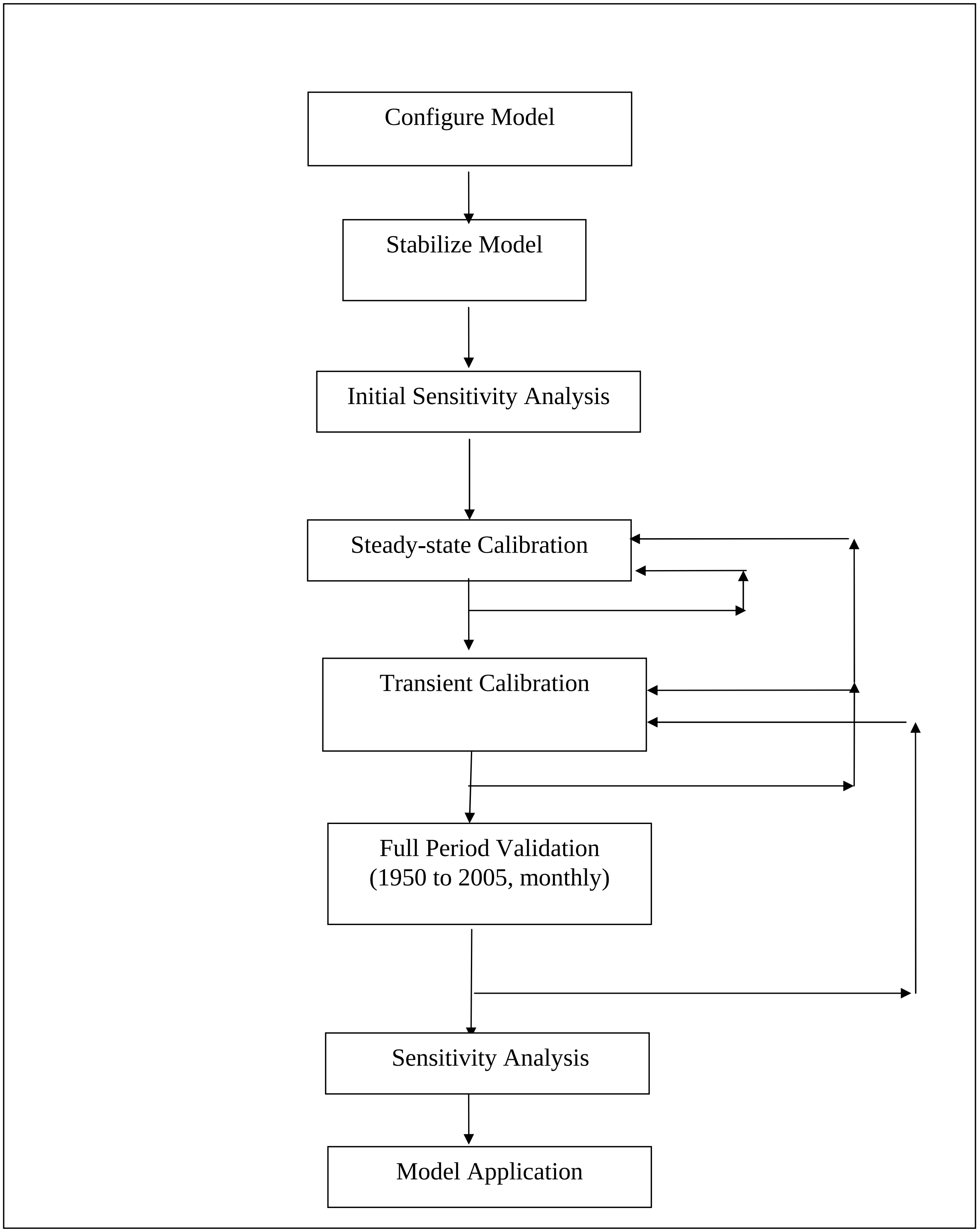
Doherty, J. 2004. PEST Model-Independent Parameter Estimation User Manual: 5th Edition. Watermark Numerical Computing. 336 pp.

Hill, M.C. and Tiedeman, C.R. 2007. Effective Groundwater Model Calibration. John Wiley & Sons, Inc. New Jersey. 455 pp.

LRE. 2006. SPDSS Task 2 – Identify Key Streamflow Gages and Estimate Streamflows for Missing Records, prepared for Colorado Water Conservation Board and Colorado Division of Water Resources, Revised August 21, 2006.

SPDSS Phase 4 Task 48

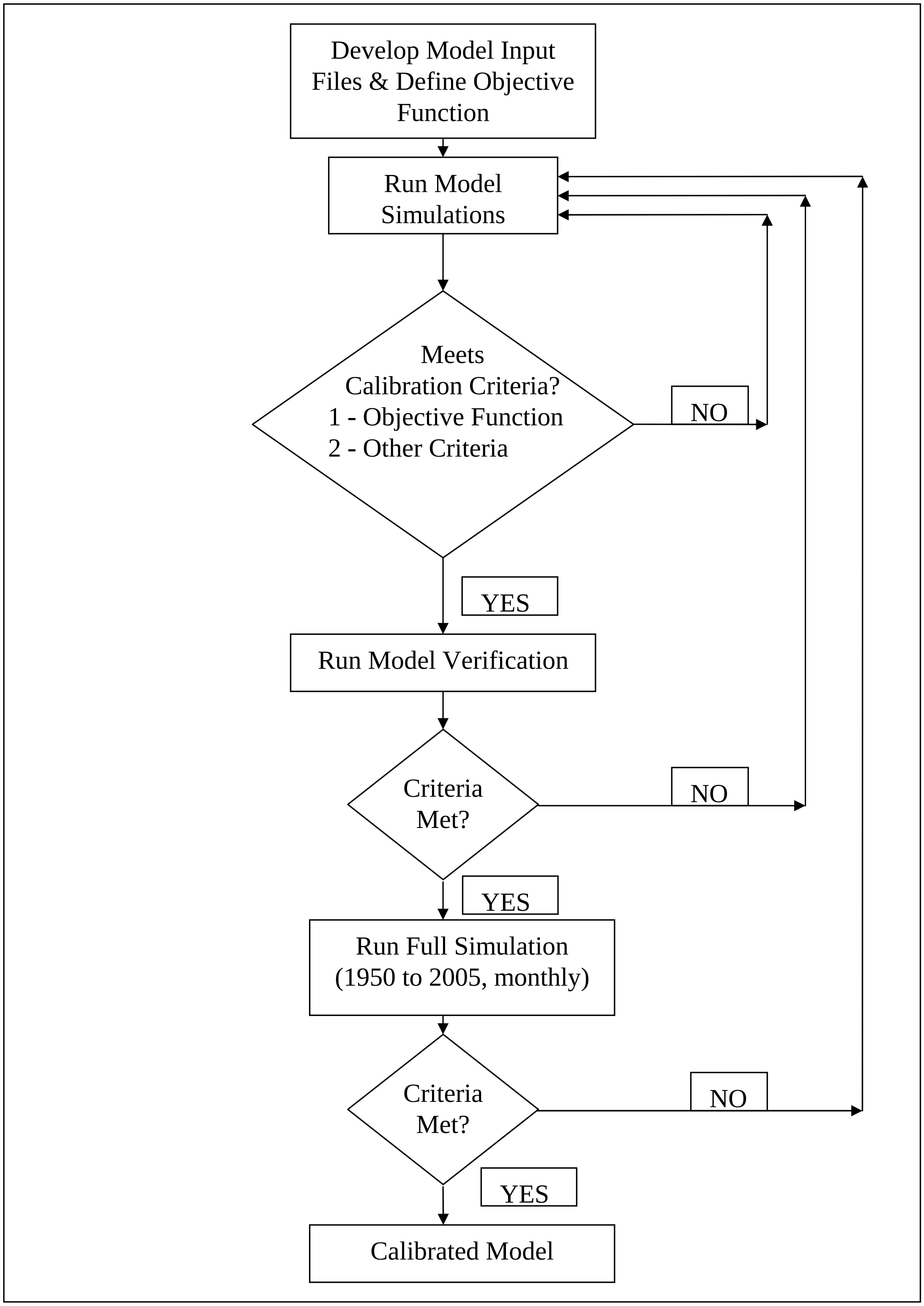
Figure 2-1. SPDSS Modeling Process



State of Colorado
Department of Natural Resources
Colorado Water Conservation Board
Division of Water Resources

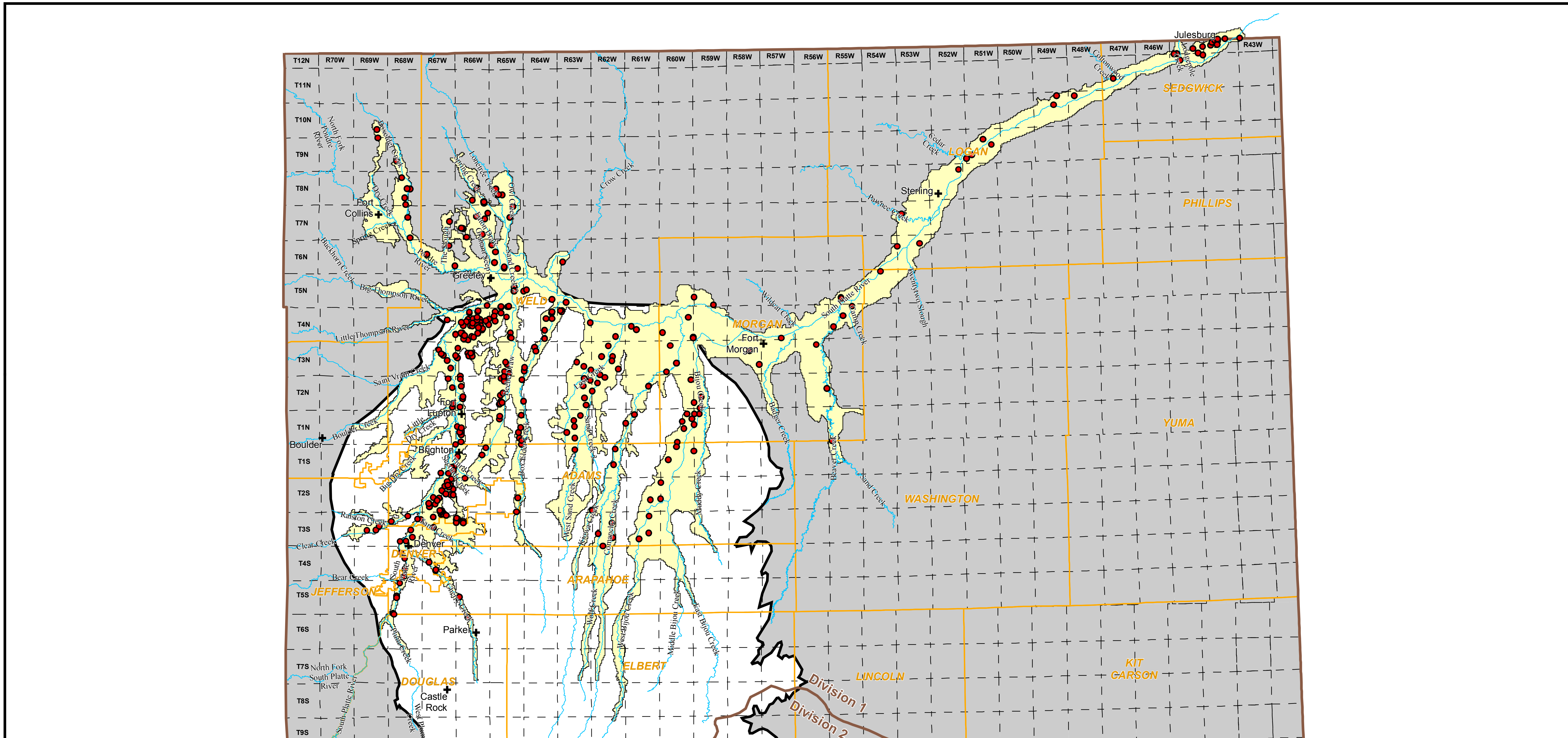
Prepared by: CDM

Figure 2-2. SPDSS Model Calibration Approach



SPDSS Phase 4 Task 48

Figure 3-1: Calibration Targets, Steady-state Period (1991 - 1994)



- Steady-state Period (January 1991 - December 1994)
- City
- Stream
- County
- Alluvial Model Extent
- Denver Basin Extent
- Township

Scale 1:1,100,000

0 5 10 20 Miles

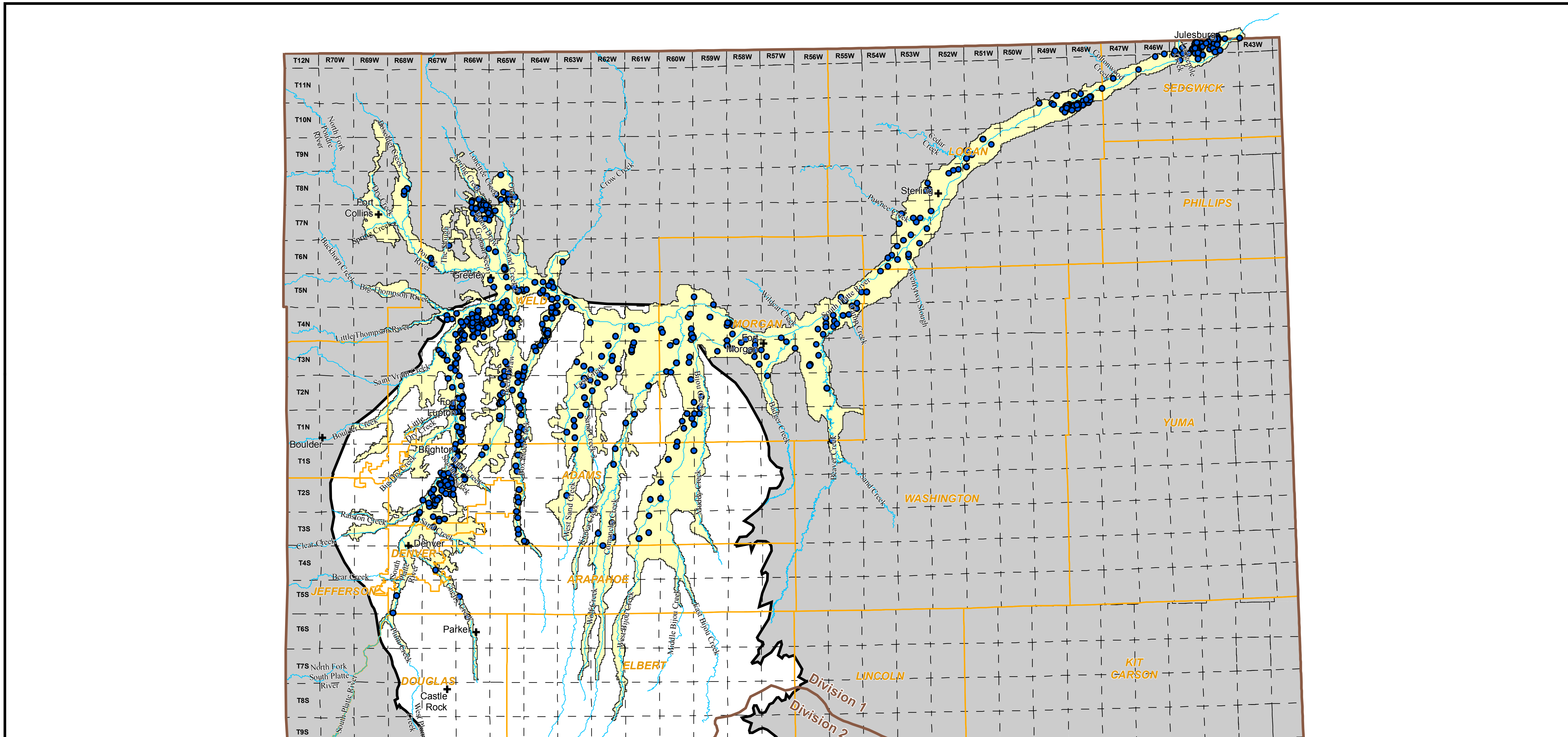
NAD 1983 UTM Zone 13N

State of Colorado
Department of Natural Resources
Colorado Water Conservation Board
Division of Water Resources

Prepared by: CDM

SPDSS Phase 4 Task 48

Figure 3-2: Calibration Targets, Transient Period (1999 - 2005)



- Transient Calibration Period (January 1999 - December 2005)
- City
- Stream
- County
- Alluvial Model Extent
- Denver Basin Extent
- Township

Scale 1:1,100,000

0 5 10 20 Miles

NAD 1983 UTM Zone 13N

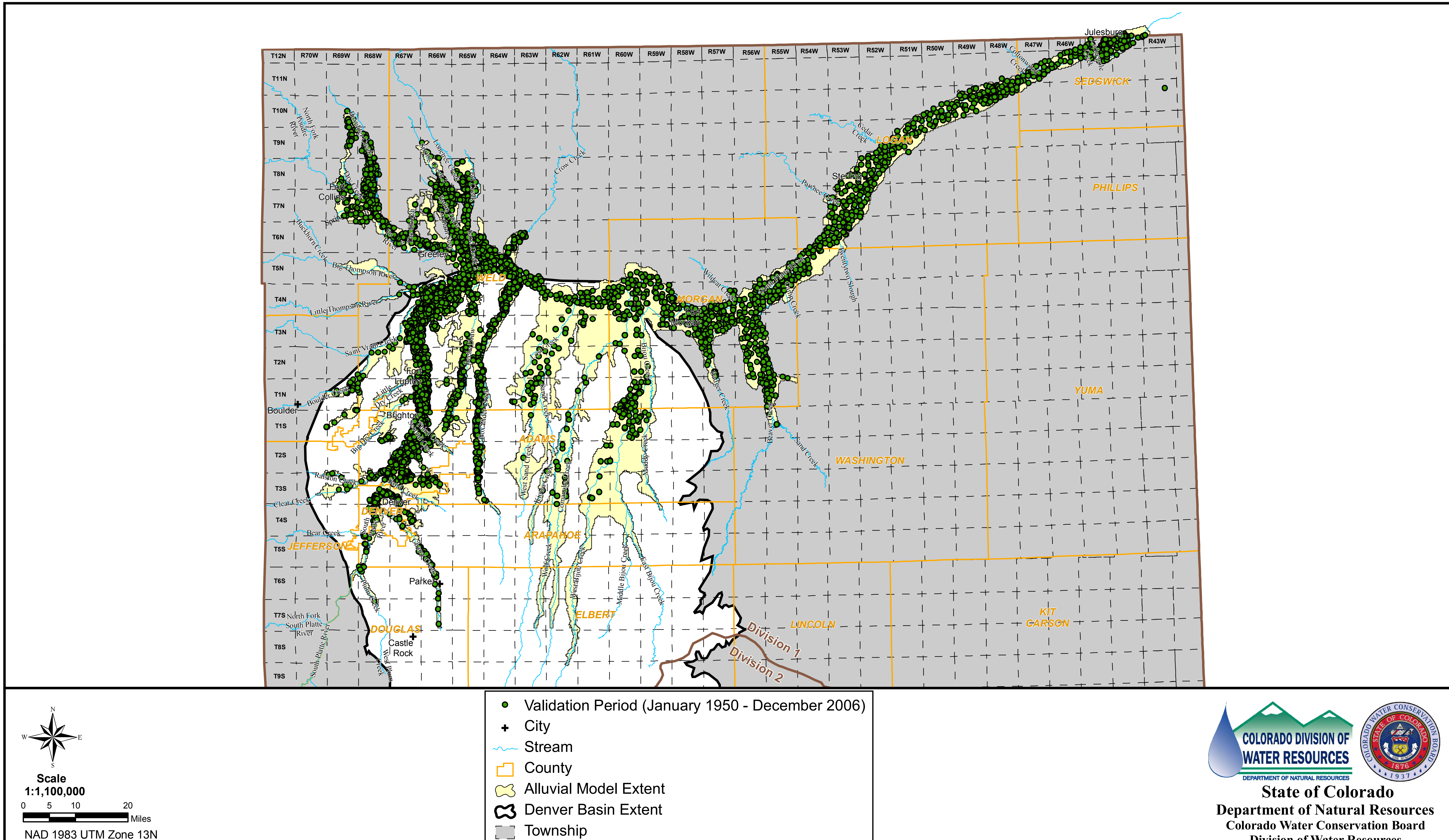
COLORADO DIVISION OF WATER RESOURCES
DEPARTMENT OF NATURAL RESOURCES

State of Colorado
Department of Natural Resources
Colorado Water Conservation Board
Division of Water Resources

Prepared by: CDM

SPDSS Phase 4 Task 48

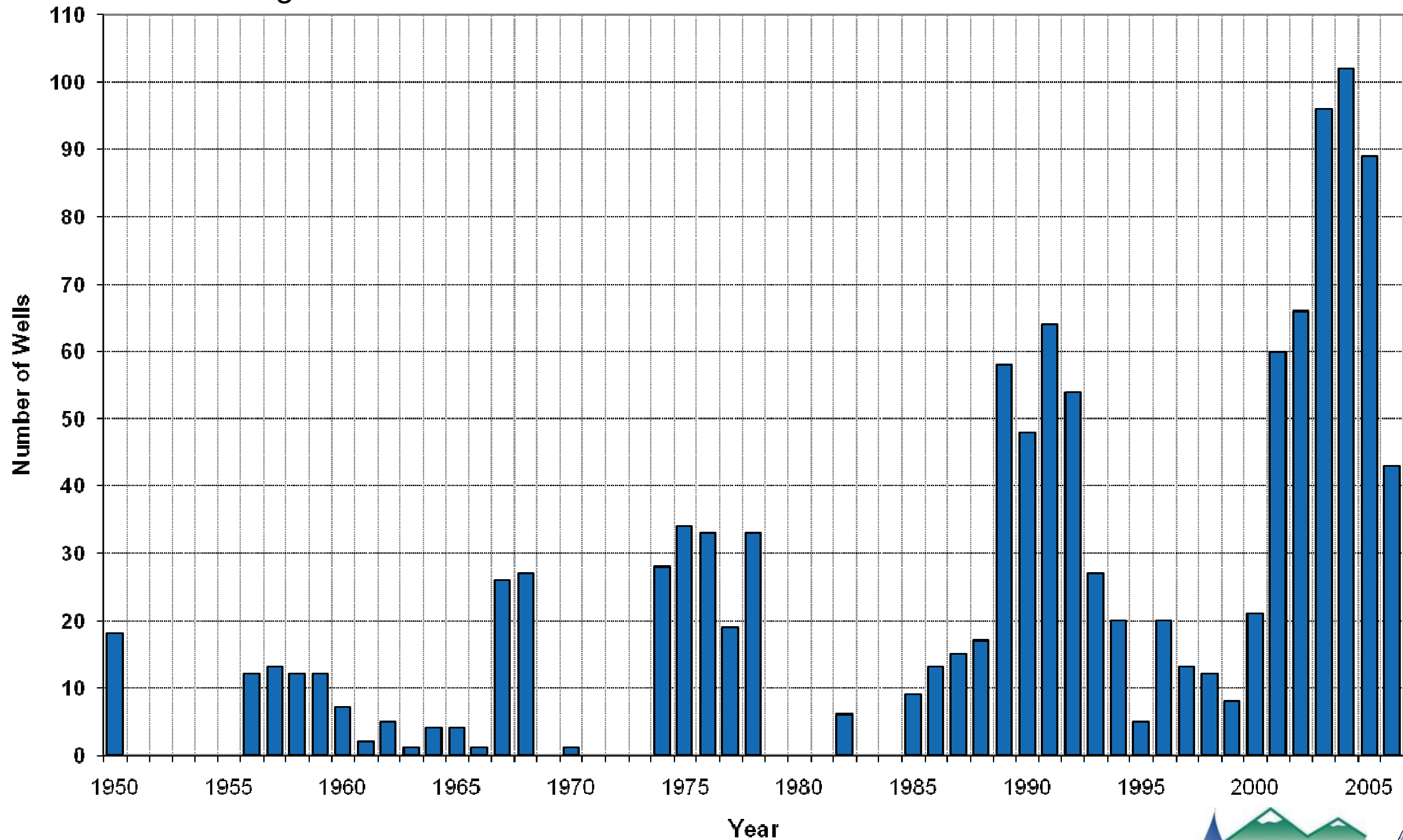
Figure 3-3: Calibration Targets, Validation Period (1950 - 2006)



Prepared by: CDM

SPDSS Phase 4 Task 48

Figure 3-4: Number of Alluvial Wells with >4 Measurements in a Year

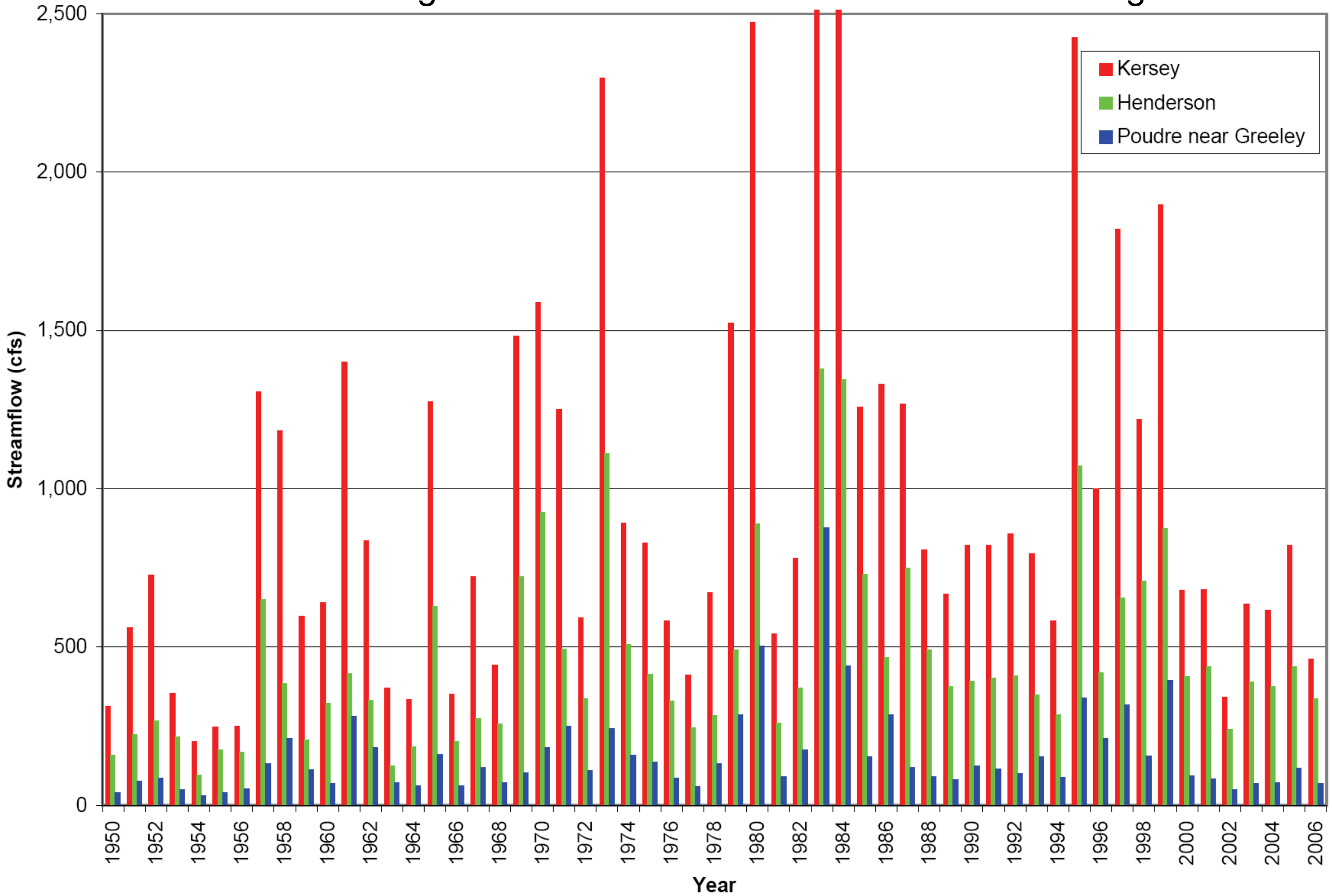


State of Colorado
Department of Natural Resources
Colorado Water Conservation Board
Division of Water Resources

Prepared by: CDM

SPDSS Phase 4 Task 48

Figure 3-5: Annual Streamflow at Selected Gages



State of Colorado

Department of Natural Resources

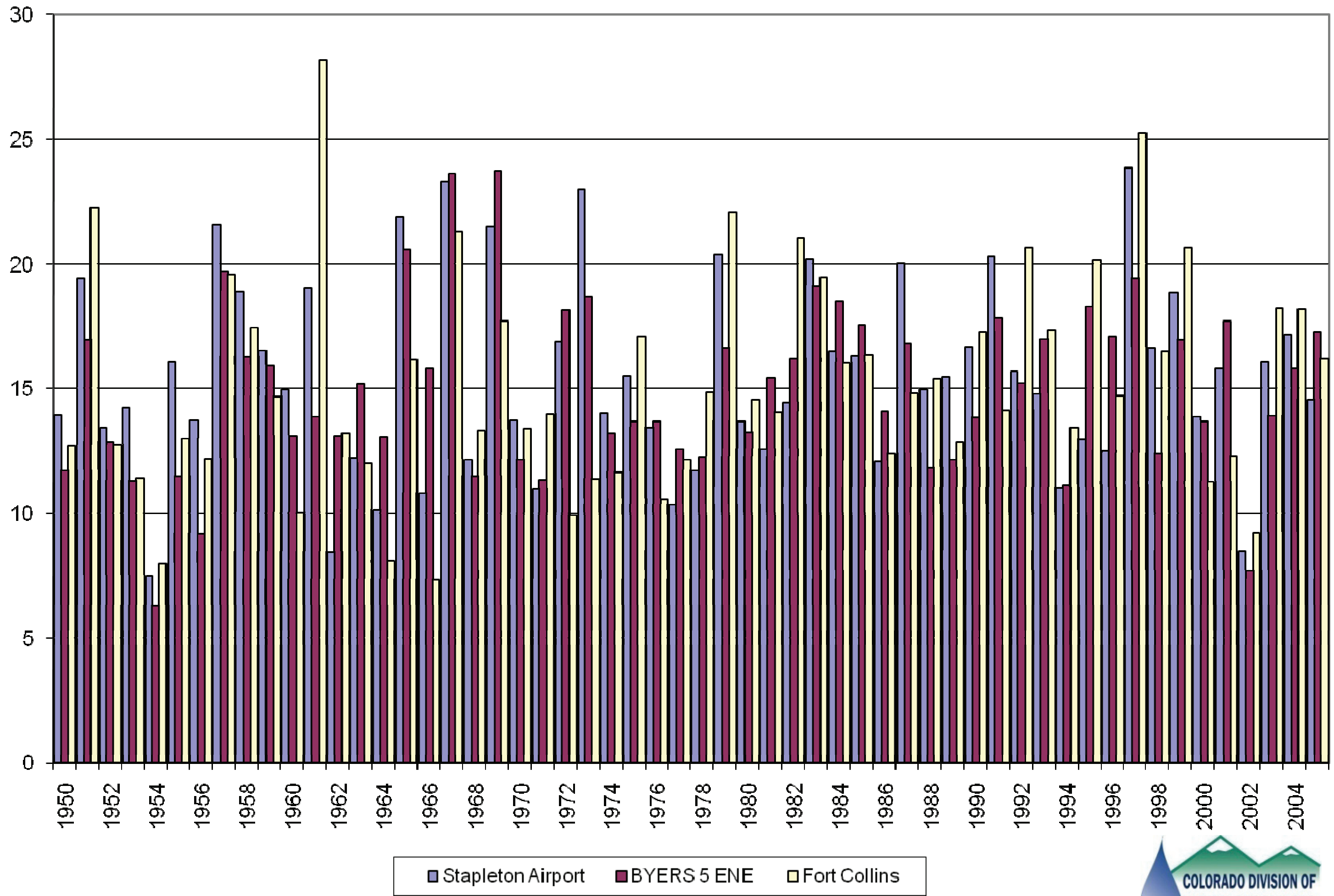
Colorado Water Conservation Board

Division of Water Resources

Prepared by: CDM

SPDSS Phase 4 Task 48

Figure 3-6: Annual Precipitation at Selected Gages (in/yr)



State of Colorado

Department of Natural Resources

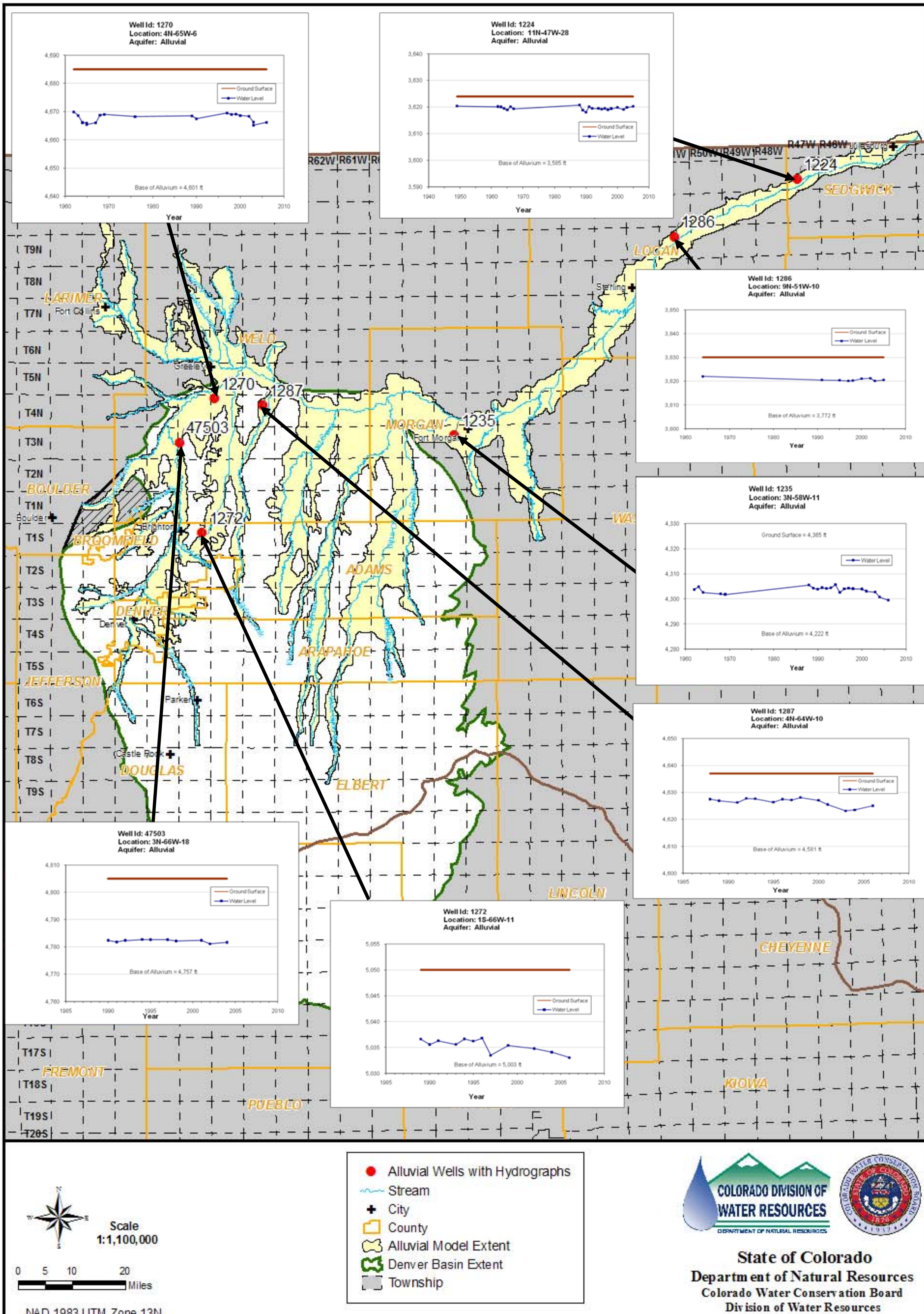
Colorado Water Conservation Board

Division of Water Resources

Prepared by: CDM

SPDSS Phase 4 Task 48

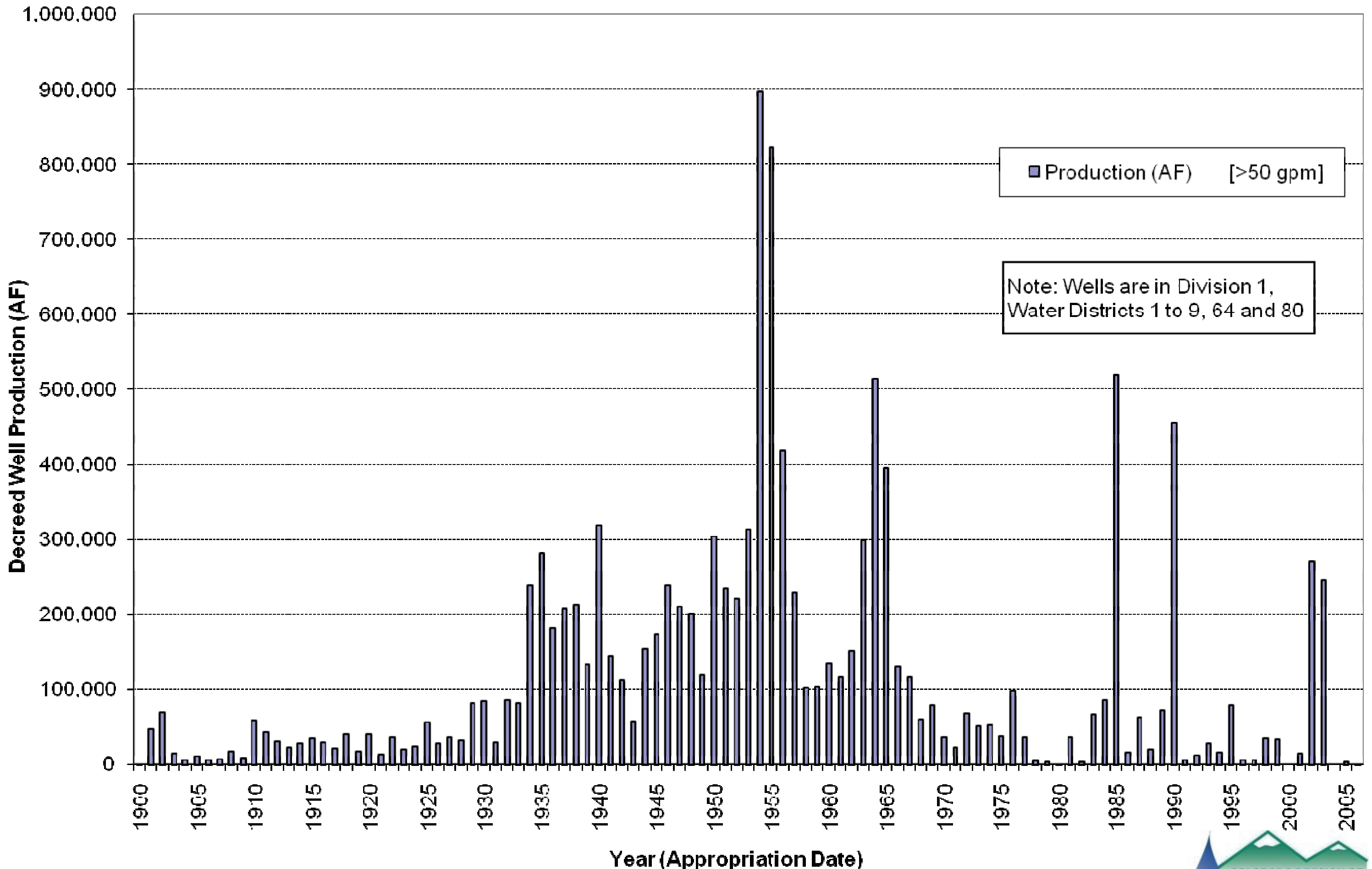
Figure 3-7: Location of Hydrographs in the South Platte Alluvial Aquifer



Prepared by: CDM

SPDSS Phase 4 Task 48

Figure 3-8: Annual Production of New Decreed Wells in the South Platte Alluvial Aquifer Model



State of Colorado
Department of Natural Resources
Colorado Water Conservation Board
Division of Water Resources

Prepared by: CDM

SPDSS Phase 4 Task 48

Figure 3-9: Calibration Targets - Streamflow Gages

State of Colorado
Department of Natural Resources
Colorado Water Conservation Board
Division of Water Resources

- Major Streamflow Gage
- City
- Stream
- County
- Active Model Area

Scale
1:1,000,000
0 5 10 20 Miles
NAD 1983 UTM Zone 13N

Prepared by: CDM