EVALUATION OF THE 1994 PREDICTIONS OF THE RUN-TIMING OF WILD MIGRANT YEARLING CHINOOK IN THE SNAKE RIVER BASIN

Prepared by:

Richard Townsend
Peter Westhagen
Dean Y asuda
John R. Skalski

Center for Quantitative Science
School of Fisheries
University of Washington
Seattle, WA

Prepared for:

U.S. Department of Energy
Bonneville Power Administration
Environment, Fish and Wildlife
P. 0. Box 3621
Portland, OR 9720x-362 1

Project Number 9 l-05 1
Task Number DE-AT79-9 1 BP 16570
Contract Number DE-AI79-87BP35885

February 1995

Table of Contents

List of Tables

List of Figures

Executive Summary

Since 1988, wild salmon have been PIT-tagged under programs conducted by the National Marine Fisheries Service, Idaho Department of Fish and Game, Oregon Department of Fish and Game and the U.S. Fish and Wildlife Service. The detection of tagged individuals at Lower Granite Darn provides a measure of the temporal and spatial distribution of the wild populations. PIT Forecaster was developed to take advantage of this historical data to predict the proportion of a particular population which had arrived at the index site in real-time and to forecast elapsed time to some future percentile in a migration.

This report evaluates the performance of the Least Squares (LS) method and the Synchronized historical pattern matching (SYNC) method used in the PIT Forecaster and two other possible methods of prediction reminiscent of methods recommended by the Fish Passage Center (FPC) (Alternatives #1 and #2). Alternative #1 bases predictions on the historical proportion of PIT-Tags recovered in a specific year and Alternative #2 uses the historical cumulative distribution of smolt predicted of a previous season. The 1994 predictions from these methods were compared to the observed distribution at the end of the 1994 season. Appendix A contains plots of the migration season comparing each predictive method to the actual proportion observed passing Lower Granite Dam.

Over the entire 1994 season, the LS method had the best prediction performance for both aggregate (mean average deviance (MAD) = 8%) and individual (MAD = 11%) streams. However. for the first half of the season, Alternative #1 was clearly a better predictor, with aggregate stream mean MADs of 2-6% and individual stream mean MADs of 4-5%. This performance deteriorated for the last half of the season, and the LS method improved to an 8% mean MAD for aggregate streams and 11% mean MAD for the individual streams. The algorithms used by PIT Forecaster determine the distribution pattern and provide a better prediction as the season progresses. Alternatives #1 and #2 depend on the year chosen as the index, without a similar correcting mechanism for a bad initial choice. For the first half of the season, though, Alternative #1 did very well for both index years selected. For the 1995 season, Alternative #1 will be combined with the LS method to create an improved predictor.

Introduction

Three Ecologically Significant Units (ESU) of Pacific salmon have been designated as either threatened or endangered (T&E) under the Endangered Species Act (ESA) in the Snake River Basin: spring/summer chinook, fall chinook and sockeye salmon. The tributary populations of spring/summer chinook reside primarily in the Salmon and Grande Ronde drainages and Imnaha River, all of which are situated upstream of Lower Granite Dam. Additionally, a small population resides in the Tucannon River, which enters the Snake River between Lower Granite and Hell's Canyon dams. The sockeye reside in the uppermost portion of the main Salmon River in the Stanley Basin. Except for the Tucannon River population, all others reside upstream from Lower Granite Dam on the Snake River.

Regulating the timing and volume of water released from storage reservoirs (flow augmentation) has become a central mitigation strategy for improving downstream migration conditions for juvenile salmonids in the Snake River. Threatened and endangered salmon stocks have received increased priority with regard to the timing of this flow augmentation. The optimum is to release water from the storage reservoirs at times when the listed stocks are in geographic locations where they encounter the augmented flow.

In the Snake River Basin, regulated water enters the system at two locations, below Hell's Canyon Dam on the Snake River and below Dworshak Dam on the Clearwater River. The preponderance of regulated water available for fish passage is provided by Dworshak Reservoir. None of the listed stocks are located in the Clearwater drainage, and thus listed stocks must migrate to below the confluence of the Clearwater and Snake Rivers before they are fully exposed to augmented flows. The confluence forms the approximate upper boundary of the Lower Granite Reservoir. Determining when stocks are in the vicinity of Lower Granite Dam and reservoir is a chief consideration in requesting flow augmentation.

Since 1988, wild salmon have been PIT-tagged under programs conducted by the National Marine Fisheries Service, Idaho Department of Fish and Game, Oregon Department of Fish and Game and the U.S. Fish and Wildlife Service. The detection of tagged individuals at Lower Granite Dam provides a measure of the temporal and spatial distribution of the wild populations. PIT Forecaster was developed to take advantage of this historical data to predict the proportion of a particular population which had arrived at the index site in real-time and to forecast elapsed time to some future percentile in a migration.

This report is a post-season analysis of the accuracy of the 1994 predictions from the PIT Forecaster program. Observed 1994 data were compared to the predictions made by PIT Forecaster for the spring outmigration of wild spring chinook observed at Lower Granite Dam through-out the season. In addition, two alternative methods were compared to PIT Forecaster. Alternative method #1 is based on the historical proportion of PIT-tags recovered for a stream for a specific year, and Alternative method #2 is based on the historical cumulative distribution of smolt passage for a specific year. These alternatives were included to give a bench-mark for the performance of the PIT Forecaster versus some possible simple methods of run estimation. Appendix A contains plots of the migration season comparing each predictive method to the actual proportion observed passing Lower Granite **Dam..**

A user's manual for the 1994 RealTime PIT Forecaster is included in Appendix B. It provides samples of the various information displays provided by the program, beyond those included in Appendix A (comparison plots of predicted versus actual season migration proportions observed). Possible displays include historical and current season plots for individual and user-defined composite streams. The PIT Forecaster's data graphical capabilities are tremendous assets in addition to its predictive functions.

Methods

Description of Data

The spring outmigration of wild spring chinook from four stream-aggregates and twelve individual streams were used in evaluating the 1994 performance of the four algorithms. The four stream aggregates include all fish tagged within a certain region, designated by its "epa-reach" (Table 1). Note that some of the individual streams were also included in the larger aggregates (Table 2). These specific aggregates and streams were chosen for their persistent recovery numbers, each having at least three years with a minimum of 30 detections per year. This was the minimum amount of historical data considered necessary in the formulation of the PIT Forecaster.

Table 1: The four stream aggregates used in evaluating the predictive performance of the PIT-Forecaster program. The individual streams which were included in the aggregates are listed below their respective aggregate.

Aggregate Name	epa-reach	Aggregate Name	epa-reach
1) Pahsimeroi River	17060202	4) Upper salmon River	17060201
-HerdCreek	17060202	-Alturas Lake Creek	17060201
- salmon River. East Fork[a]	17060202	-FourthofJulyCreek	17060201
2) Salmon River, Upper Middle Fork	17060205	-Frenchman Creek	17060201
- Bear Valley Creek[a]	17060205	-HuckleberryCreek	17060201
- Capehorn Creek	17060205	-Pettit Lake Creek	17060201
-Elkcreek	17060205	-Pole Creek	17060201
-LoonCreek	17060205	-Red Fish Lake Creek	17060201
-MarshCreek	17060205	- Smiley Creek	17060201
3) Salmon River south Fork	17060208	-Valley Creek[a]	17060201
- Johnson Creek	17060208		
-LakeCreek	17060208		
- Salmon River South Fork[a]	17060208		
-SeceshRiver[a]	17060208		

a Also used in separate stream analysis.

Table 2: The twelve individual rivers used in evaluating the predictive performance of the PIT-Forecaster program.

Stream Name	epa-reach
Bear Valley Creek	17060205
Big Creek	17060206
Catherine Creek	17060104
Elk creek	17060205
Imnaha River	17060102
Lostine River	17060105
Marsh Creek	17060205
Salmon River	17060209
Salmon River. East Fork	17060202
Salmon River South Fork	17060208
Secesh River	17060208
valley creek	17060201

Prediction Models

The following is a brief synopsis of the four methods used in this report. PIT Forecaster has two methods of prediction. A least-squares (LS) method and a synchronized historical pattern-matching (SYNC) method. The two other methods are based on a simple assumption that the run for the present year is directly related to a previous outmigration pattern.

Least-Squares Method

For a given day in the run, the LS algorithm computes the predicted percentage ($\hat{p}$) of the outmigration by finding the value of $\hat{p}$ that minimizes the squared error according to historical runs. The squared error (SS) for each $\hat{p}$ is summed over the historical years for which data are available. Each outmigration pattern is divided into 100 equal portions and the slopes at each corresponding point are computed. The sum of squares for a prediction compares the slopes for the current year (s_{oj}) versus the respective slopes for the historical years (s_{ij}). The total squared error for each predicted percentage of outmigration $\hat{p}$ is calculated according to the formula

$$SS(\hat{p}) = \sum_{i=1}^{n} \sum_{j=0}^{p} (s_{oj} - s_{ij})^2 w_{ij} \tag{1}$$

where s_{oj} = observed slope at the jth percentile (j = 0, ... p) for the current year of prediction,
s_{ij} = slope at the jth percentile (j = 0, ... p) for the ith historical year (i = 1, ... n), and
w_{ij} = weight for the jth percentile for the ith historical year.

For example, letting $\hat{p}$ = 30%, the present run will be compared to the first 30% of the outmigration for each historical year. Similar calculations are performed for each percentage from 0 to 100 percent. The percentage that minimizes the sum of squares (1) is the best prediction for the current outmigration timing. The weighting factor is included to more evenly distribute the squared error contribution throughout the outmigration distribution. The weights are

$$w_{ij} = \frac{D_{oi} + D_{ii}}{R_o + R_i} \tag{2}$$

where D_{oj} = estimated number of days between the (j- 1) and jth percentile for the present year,
D_{ij} = number of days between the (j- 1) and j^{th} percentile for the ith historical year (i = 1, ..., n),
R_o = range in days of the current observed outmigration, and
R_i = range in days of the ith historical year outmigration (i = 1, ... n).

The effect of w_{ij} is to give more weight to the errors generated from the tails of the distribution, where the slopes tend to be flat and the number of days between each percentile point are high. Less weight is given to the mid-season, where large slopes are more likely. The total sum of the weights adds to one.

Synchronized Historical Pattern Matching

The SYNC method selects the $\hat{p}$ which minimizes the error between the present year and a historical-year aggregate. The algorithm finds the first significant peak in the density of each historical run and then aligns each of the densities on these peaks To determine the peak, the local minima[1] in the first half of the historical-run density are determined. The area between each of these minima is calculated, and the significant peak lies within the area where i is minimized (i.e. the earliest qualifying occurrence in the outmigration pattern) in the equation

$$A_{(i-1,\, i)} > 0.2B_{i-1} \tag{3}$$

where $A_{(i\text{-}1,\, i)}$ = the area under the historical-run density between the (i- 1) and ith local minima (i = 1, ... n), and
$B_{i\text{-}1}$ = the area under the historical-run density between the (i-1)th local minima and the 50th percentile.

The local maximum between the local minima determined by (3) is the significant peak the historical years are aligned on. The average of these aligned historical densities is the historical aggregate density. The squared error for each $\hat{p}$ is calculated according to the formula

$$\text{SS } (\hat{p}_i) = \sum_{i=O}^{\hat{p}} (S_{oi} - S_{ai})^2 \tag{4}$$

where S_{oi} = slope at the i^{th} percentile (i = 0, ... $\hat{p}$) in the current year outmigration density, and
S_{ai} = slope at the i^{th} percentile (i = 0, ... $\hat{p}$) in the aggregate outmigration density.

1. A minima is a point on the density line where the first derivative of the density at that point equals zero and the second derivative is positive.

Alternative #1

Alternative #1 makes a prediction of run timing by using the total recapture proportion observed in a previous season and then assuming that proportion to be consistent for the present year. The predicted percent of the run is calculated according to the formula

$$\hat{p} = \frac{x_d}{p_y \times N} \tag{5}$$

where $\hat{p}$ = the estimated proportion of the outmigration passed on day *d*,
x_d = the total observed smolt to day *d*,
p_y = the total proportion of outmigration observed for historical year y, and
N = the total number of smolt tagged for the present year.

The number of fish tagged for the present year for a given stream or stream aggregate is multiplied by the recapture ratio (p_y) of a previous year (Table 3) to determine the total number of fish expected. The proportion passed is then estimated. For example, the Pahsimeroi River observed a recapture ratio of 5.1% in 1992 and 4.3% in 1993. The expected total number of smolt for 1994 based on 1992 data would be 133.67 smolt (2621 x 0.051) or, based on 1993, 112.7 (2621 x 0.043) smolt. The observed recovery proportions for each of the 1992 and the 1993 seasons were used to make predictions for the 1994 season.

Alternative #2

In method #2 the prediction was the historical proportion observed on a given day of outmigration for a specified historical year.

$$\hat{p} = p_{yd} \tag{6}$$

where $\hat{p}$ = the estimated proportion of the outmigration observed on day *d*, and
p_{yd} = the proportion of outmigration passed on day *d* for historical year y.

Both 1992 and 1993 recapture proportions were used to make prediction for 1994. For a given day of run, the proportion predicted is given by the proportion observed in the index year on that day of the run (i.e. for a run estimated to be in its 15th day, the percentage passed by day 15 in the historical run of 1992 or 1993 is the estimated present percentage observed).

Calculation of Comparison Scores

The results presented in Tables 4 through 9 contain the mean absolute deviance (MAD) of the predictions from the observed 1994 data for each stream and stream aggregate. The MAD is calculated by the formula

$$\text{MAD} = \frac{\sum_{i=1}^{n} \left| \hat{p}_i - p_i \right|}{n} \qquad (7)$$

where $\hat{p}_i$ = predicted cumulative percentage of run completed for day **i**,
p_i = observed cumulative percentage of run completed for day i, and
n = total number of days in run for 1994 season.

The methods are compared three ways: the MAD over the entire run, the MAD over the first half of the run (i.e. cumulative run to the 50%). and the MAD over the last half of the run.

Table *3:* **The Number of smolt released and detected at Lower Granite dam. The proportion ($\hat{p}_v$) is the number of detected smolt divided by the total number tagged.**

Tagging Site	1992			1993			1994		
	# Tagged	# Detected	Proportion $\hat{p}_v$	# Tagged	# Detected	Proportion $\hat{p}_v$	# Tagged	# Detected	Proportion $\hat{p}_v$
Aggregate Streams									
Pahsimeroi River	979	50	0.051	2365	101	0.043	2621	145	0.055
Salmon River Upper Middle Fork	2694	191	0.071	2797	330	0.118	10180	541	0.053
Salmon River South Fork	2039	121	0.059	3067	343	0.112	6559	481	0.073
Upper Salmon River	2303	57	0.025	4122	140	0.034	2515	59	0.023
Individual Streams									
Bear Valley Creek	1042	69	0.066	1015	79	0.078	934	63	0.067
Big Creek	1002	57	0.057	733	70	0.095	721	40	0.055
Catherine Creek	940	67	0.071	1095	119	0.109	1000	77	0.077
Elk Creek	462	36	0.078	628	48	0.076	998	54	0.054
Imnaha River	815	94	0.115	1237	180	0.146	1747	1%	0.112
Lostine River	1107	92	0.083	1000	132	0.132	725	67	0.092
Marsh Creek	981	67	0.068	949	139	0.146	3249	393	0.121
Salmon River	23%	65	0.027	1498	40	0.027	736	40	0.054
Salmon River East Fork	669	33	0.049	843	40	0.047	883	45	0.051
Salmon River South Fork	1027	81	0.079	2106	245	0.116	5885	442	0.075
Secesh River	1012	40	0.040	327	31	0.095	422	39	0.092
Valley Creek	969	34	0.035	1026	36	0.035	848	39	0.046

Results

Comparing the performance of the prediction methods across the entire season indicates that, on average, the LS procedure performed the best for 1994 (Tables 4-5). For the four aggregate streams (Table 4), LS had the smallest mean MAD (8%) and median MAD (6%) of all methods. The SYNC algorithm performed slightly worse with a mean MAD of 11%. The two simplified algorithms varied in performance depending on whether 1992 or 1993 was used as the index year (mean MAD 12-26% for Alternative #1 10-15% for Alternative #2). The LS algorithm also was the best predictor with a mean MAD of 11% for the individual streams (Table 5). Alternative #2 had nearly as good of a performance with mean MAD of 1 1-13%, again depending on the selected index year. However, the performance of the different predictors is somewhat obscured by examination of the average performance across the entire season. Additional information on predictive performance can be acquired by examining the early and latter halves of the season.

Table 4: Comparison of mean absolute deviances (MAD) for selected stream aggregates for the entire 1994 outmigration.

Aggregate	PIT Forecaster		Alternative #1		Alternative #2	
	LS	SYNC	1992	1993	1992	1993
Pahsimeroi River	0.06	0.08	0.10	0.19	0.02	0.12
Salmon River Upper Middle FOrk	0.07	0.14	0.18	0.41	0.13	0.11
Salmon River South Fork	0.05	0.08	0.19	0.25	0.08	0.07
Upper Salmon River	0.14	0.15	0.01	0.20	0.17	0.28
mean MAD	0.08	0.11	0.12	0.26	0.10	0.15
median MAD	0.06	0.11	0.14	0.22	0.11	0.12
range	.05-.14	.08-.15	.01-.19	.19-.41	.02-.17	.07-.28

Table 5: Comparison of mean absolute deviances (MAD) for selected single streams for the entire 1994 outmigration.

Tagging Site	PIT Forecaster		Alternative #l		Alternative #2	
	LS	SYNC	1992	1993	1992	1993
Bear Valley Creek	0.08	0.16	0.05	0.08	0.11	0.14
Big Creek	0.20	0.23	0.01	0.29	0.22	0.06
Catherine Creek	0.06	0.06	0.13	0.20	0.03	0.06
Elk Creek	0.08	0.14	0.19	0.21	0.19	0.19
lmnaha River	0.11	0.13	0.01	0.17	0.08	0.03
Lostine River	0.06	0.06	0.12	0.20	0.06	0.07
Marsh Creek	0.07	0.13	059	0.14	0.12	0.08
Salmon River	0.30	0.33	0.78	0.73	0.26	024
Salmon River East Fork	0.10	0.13	0.06	0.04	0.11	0.08
Salmon River South Fork	0.05	0.08	0.02	0.27	0.05	0.06
Secesh River	0.17	0.14	1.17	0.04	0.21	0.15
Valley Creek	0.06	0.14	026	0.18	0.08	0.11
mean MAD	0.11	0.14	028	0.21	0.13	0.11
median MAD	0.08	0.14	0.12	0.19	0.11	0.08
range	.05-.30	.06-.33	.0l-1.17	.04-.73	.03-26	M-24

Tables 6 and 7 examine the performances of the methods during the first half of the outmigration. For the aggregate streams (Table 6), Alternatives #l and #2 both outperform the methods of the PIT-Forecaster with mean MADs of 2-6% and 6- 11% respectively. Similarly, for the individual streams (Table 7), Alternatives #l and #2 were better than both LS and SYNC, regardless of the historical year used as the index year. The mean MAD for Alternative #l was 4-5%. The better performance of either alternative method can be attributed to the fact that there is little historical pattern to lock the LS or SYNC algorithms on at the start of the season. Historical recovery rates provide better priors for prediction early into the outmigration.

Table 6: Comparison of mean absolute deviances for selected stream aggregates for the first half of the observed 1994 outmigration.

Aggregate	PIT Forecaster		Alternative #l		Alternative #2	
	LS	Sync	1992	1993	1992	1993
Pahsimeroi River	0.11	0.11	0.02	0.04	0.02	0.10
Salmon River upper Middle Fork	0.12	0.10	0.03	0.06	0.06	0.07
Salmon River South Fork	0.03	0.02	0.05	0.07	0.07	0.08
Upper Salmon River	0.17	0.18	0.00	0.06	0.11	0.19
meanMAD	0.11	0.10	0.02	0.06	0.06	0.11
median MAD	0.12	0.11	0.02	0.06	0.06	0.09
range	.03-.17	.02-.18	.00-.05	.04-.07	.02-.ll	.08-.19

Table 7: Comparison of mean absolute deviances for selected single streams for the first half of the observed 1994 outmigration.

Tagging Site	PIT Forecaster		Alternative #l		Alternative #2	
	LS	Sync	1992	1993	1992	1993
Bear Valley Creek	0.15	0.13	0.01	0.02	0.08	0.08
Big Creek	0.14	0.13	0.00	0.06	0.10	0.03
Catherine Creek	0.09	0.10	0.03	0.04	0.08	0.11
Elk Creek	0.15	0.11	0.04	0.04	0.09	0.12
lmnaha River	0.18	0.19	0.00	0.01	0.16	0.07
Lostine River	0.14	0.11	0.03	0.05	0.07	0.09
Marsh Creek	0.11	0.09	0.09	0.02	0.06	0.06
Salmon River	0.21	0.20	0.12	0.11	0.10	0.10
Salmon River East Fork	0.09	0.10	0.01	0.01	0.04	0.08
Salmon River South Fork	0.04	0.04	0.01	0.07	0.04	0.08
Secesh River	0.13	0.12	0.18	0.01	0.11	0.08
Valley Creek	0.16	0.16	0.08	0.05	0.07	0.15
mean MAD	0.13	0.12	0.05	0.04	0.08	0.09
median MAD	0.14	0.12	0.03	0.04	0.08	0.08
range	.04-.21	.04-.20	.00-.18	.0l-.1l	.04-.16	.03-.15

Tables **8-9** summarize performances during the second half of the outmigration. The performance of the algorithms is reverse that observed during the first half of the season. The LS and SYNC methods performed better than Alternatives #l and #2. The LS method has the smallest mean MAD **(8%)** and median MAD (6%) of all methods for the aggregate streams. For individual streams, LS again had the smallest median MAD (6%) and among the smallest mean MAD (11%). The performance of the LS method improved as the season progressed while the error of the simplified methods increased. Appendix A presents plots of the progress of the predictions over the season for all streams and aggregates. These plots show that the LS and SYNC methods converged on the true outmigration patterns as the season progressed. The simplified alternatives lack this property, accounting for their poor performance in the latter half of the season.

Table 8: Comparison of mean absolute deviances (MAD) for selected stream aggregates for the last half of observed 1994 outmigration.

Aggregate	PIT-Forecaster		Alternative #l		Alternative #2	
	LS	Sync	1992	1993	1992	1993
Pahsimeroi River	0.05	0.07	0.12	022	0.03	0.12
Salmon River Upper Middle Fork	0.06	0.15	022	0.50	0.14	0.13
Salmon River South Fork	0.06	0.10	0.24	032	0.09	0.06
Upper Salmon River	0.13	0.15	0.01	024	0.19	0.30
mean MAD	0.08	0.12	0.15	0.32	0.11	0.15
median MAD	0.06	0.12	0.17	028	0.12	0.12
range	.05-.13	.07-.15	.0l-.24	.22-.50	.03-. 19	.06-.30

Table 9: Comparison of mean absolute deviances (MAD) for selected single streams for the last half of the observed 1994 outmigration.

Tagging Site	PIT Forecaster		Alternative #l		Alternative #2	
	LS	Sync	1992	1993	1992	1993
Bear Valley Creek	0.06	0.17	0.07	0.11	0.12	0.16
Big Creek	0.21	0.25	0.02	0.36	025	0.07
catherine Creek	0.05	0.05	0.16	0.25	0.02	0.04
Elk Creek	0.07	0.15	024	026	021	0.21
Imnaha River	0.10	0.11	0.01	023	0.06	0.02
Lostine River	0.04	0.05	0.15	0.25	0.06	0.07
Marsh Creek	0.06	0.13	0.74	0.18	0.13	0.08
salmon River	0.31	0.35	0.97	0.90	0.31	0.28
Salmon River East Fork	0.11	0.14	0.08	0.05	0.13	0.08
Salmon River South Fork	0.05	0.10	0.03	0.34	0.05	0.06
Secesh River	0.17	0.14	1.31	0.05	022	0.16
Valley Creek	0.05	0.14	0.32	022	0.08	0.10
mean MAD	0.11	0.15	0.34	027	0.14	0.11
median MAD	0.06	0.14	0.16	024	0.12	0.08
range	.04-.3 1	.05-.35	.01-l.31	.05-.90	.02-.31	.02-.28

Discussion

The results from the 1994 season are encouraging and suggest improvements that can be made to the PIT Forecaster program. The LS method had an average MAD of only 8% over the entire 1994 observed season. Although not the best method for the start of the season, it improved markedly by the end of the season. Alternative #l's performance in the first half of the season recommends its incorporation into the PIT Forecaster to improve the performance of the program. The predictions by the two methods will be weighted through the season, giving more weight to predictions made by Alternative #l at the beginning and less as the season progresses, where the LS method proved to be more accurate.

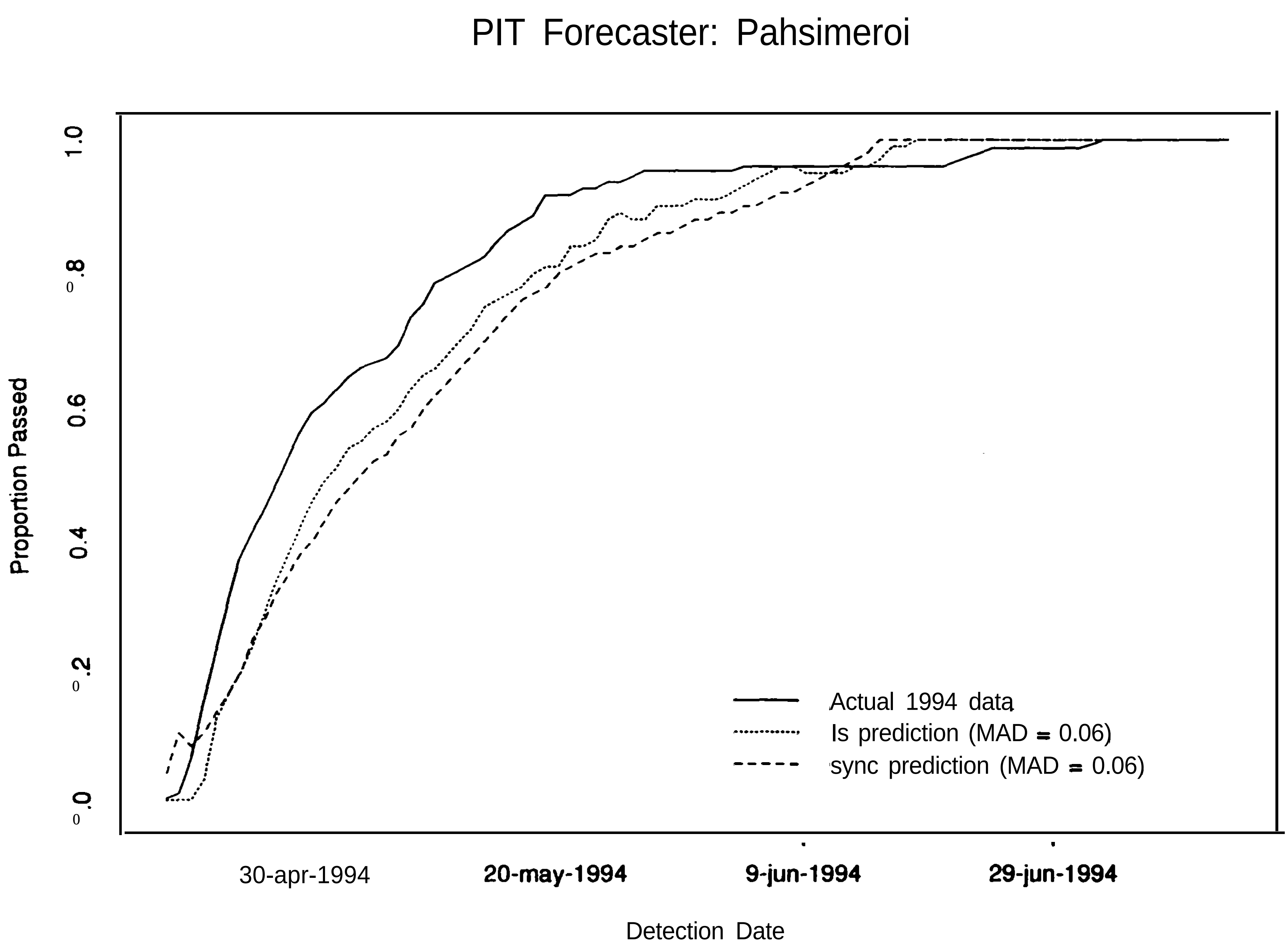
The second change for 1995 will be the removal of the SYNC method from the program. Its original inclusion into the PIT Forecaster was as a time/resource-saving method which would give comparable results without the computational intensity of the LS method. As present computer resources have been quite adequate, the removal of the less accurate SYNC method will remove conflicting recommendations from the program.

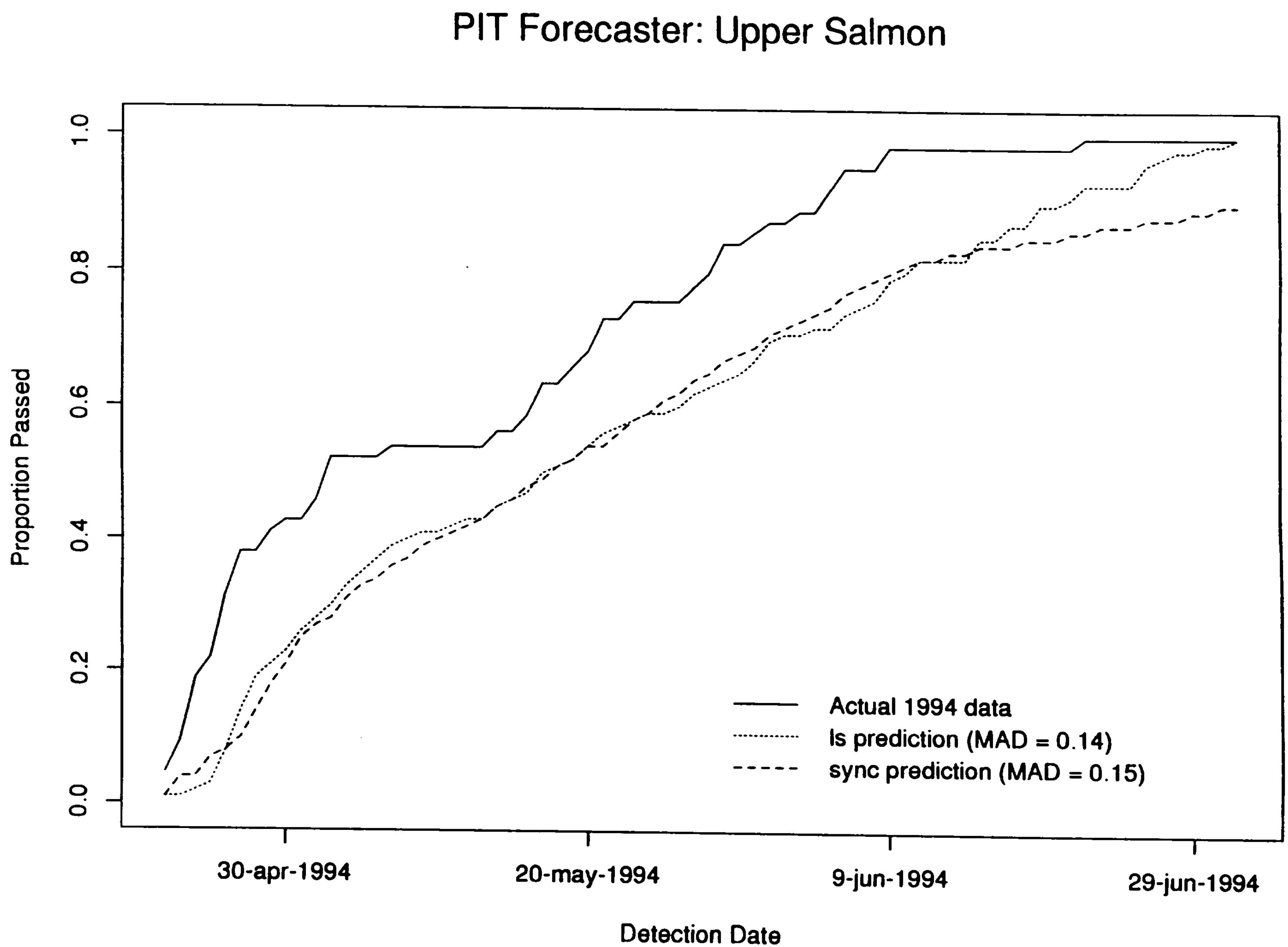
In this analysis, the streams that were included in an aggregate were fixed. A third change will allow the user to determine the individual streams to be included into a composite for outmigration prediction. This enables the user to tailor the predictions to specific areas of interest. Furthermore, rather than predicting an average outmigration for the streams within an aggregate, the new method will weight predictions on the expected numbers of PIT-tagged salmon outmigrating from their respective streams.

Appendix A

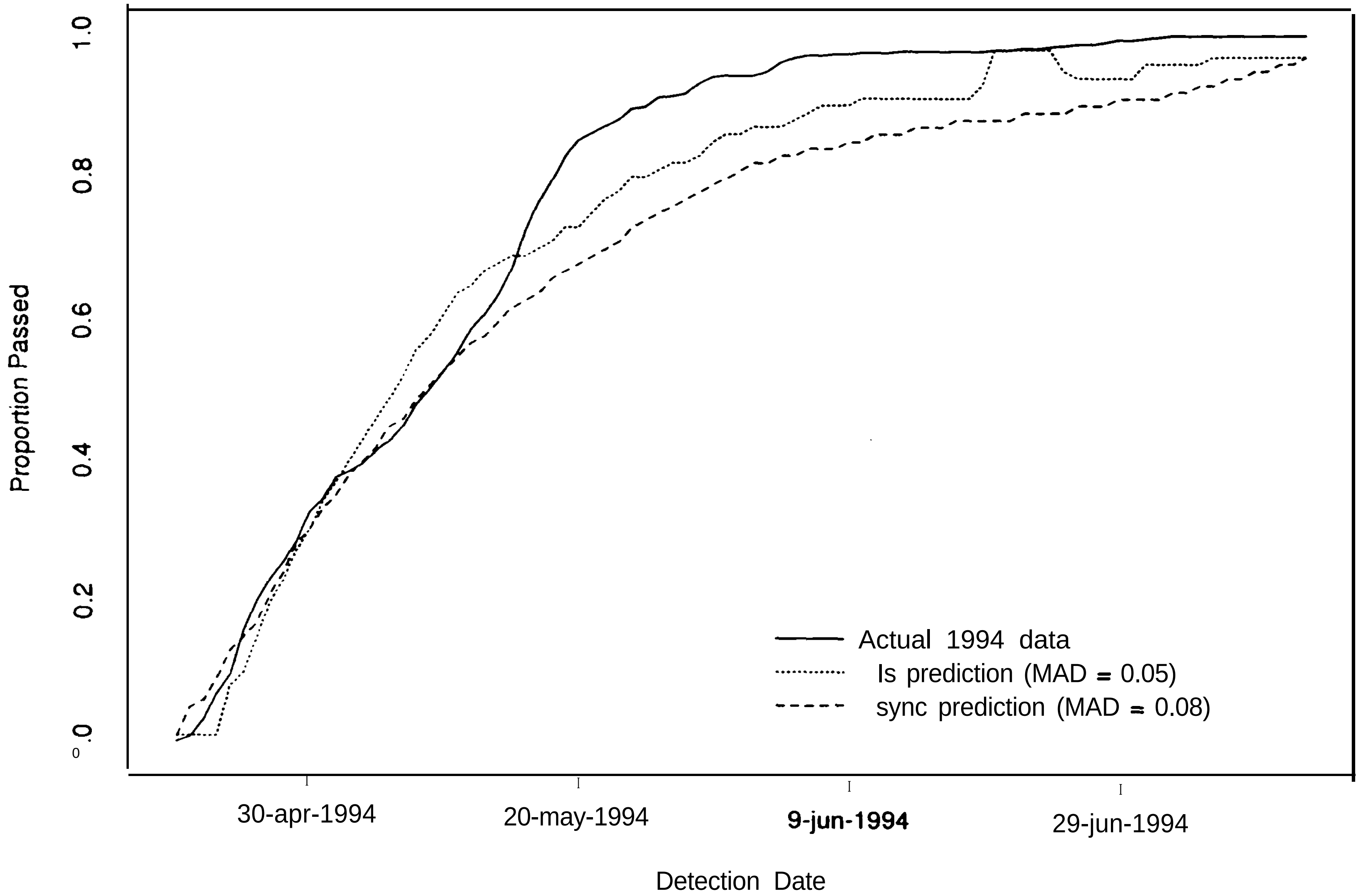
Performance Plots for the 1994 Outmigration Season

PIT Forecaster Plots

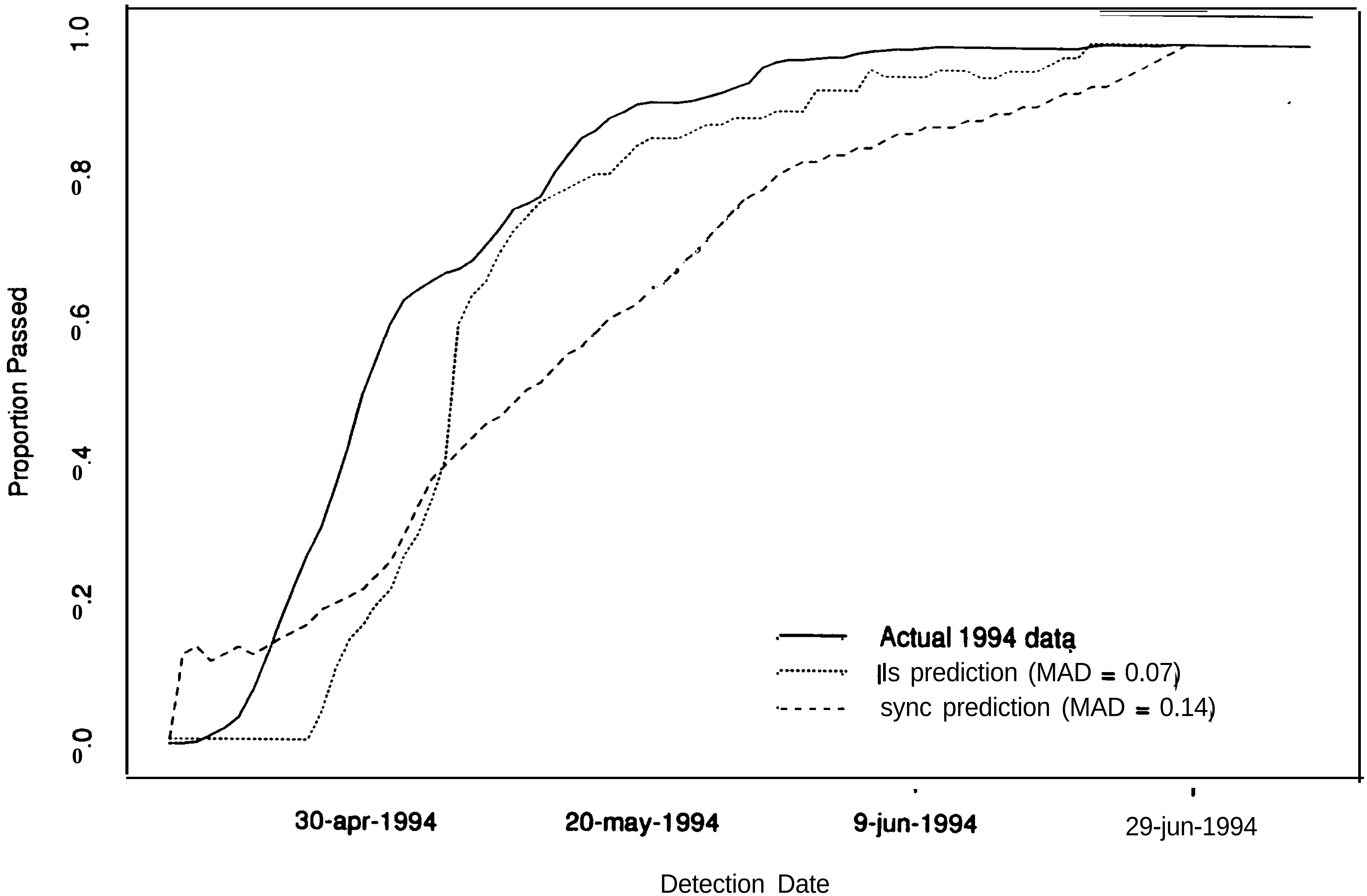

PIT Forecaster: Pahsimeroi
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
Actual 1994 data
ls prediction (MAD = 0.06)
sync prediction (MAD = 0.06)
30-apr-1994
20-may-1994
9-jun-1994
29-jun-1994
Detection Date

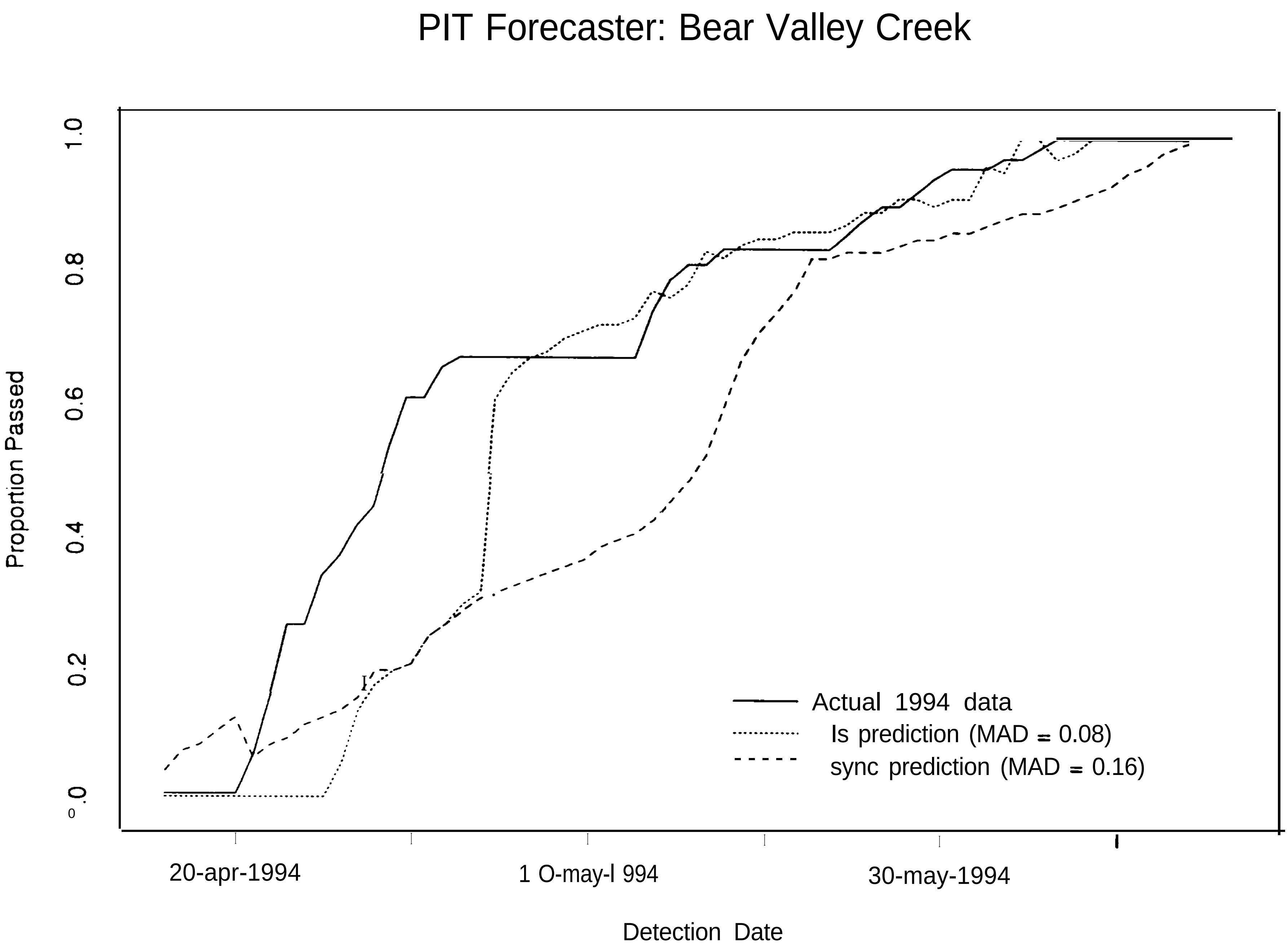

PIT Forecaster: Upper Salmon
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
30-apr-1994
20-may-1994
9-jun-1994
29-jun-1994
Detection Date
Actual 1994 data
ls prediction (MAD = 0.14)
sync prediction (MAD = 0.15)

PIT Forecaster: South Fork Salmon

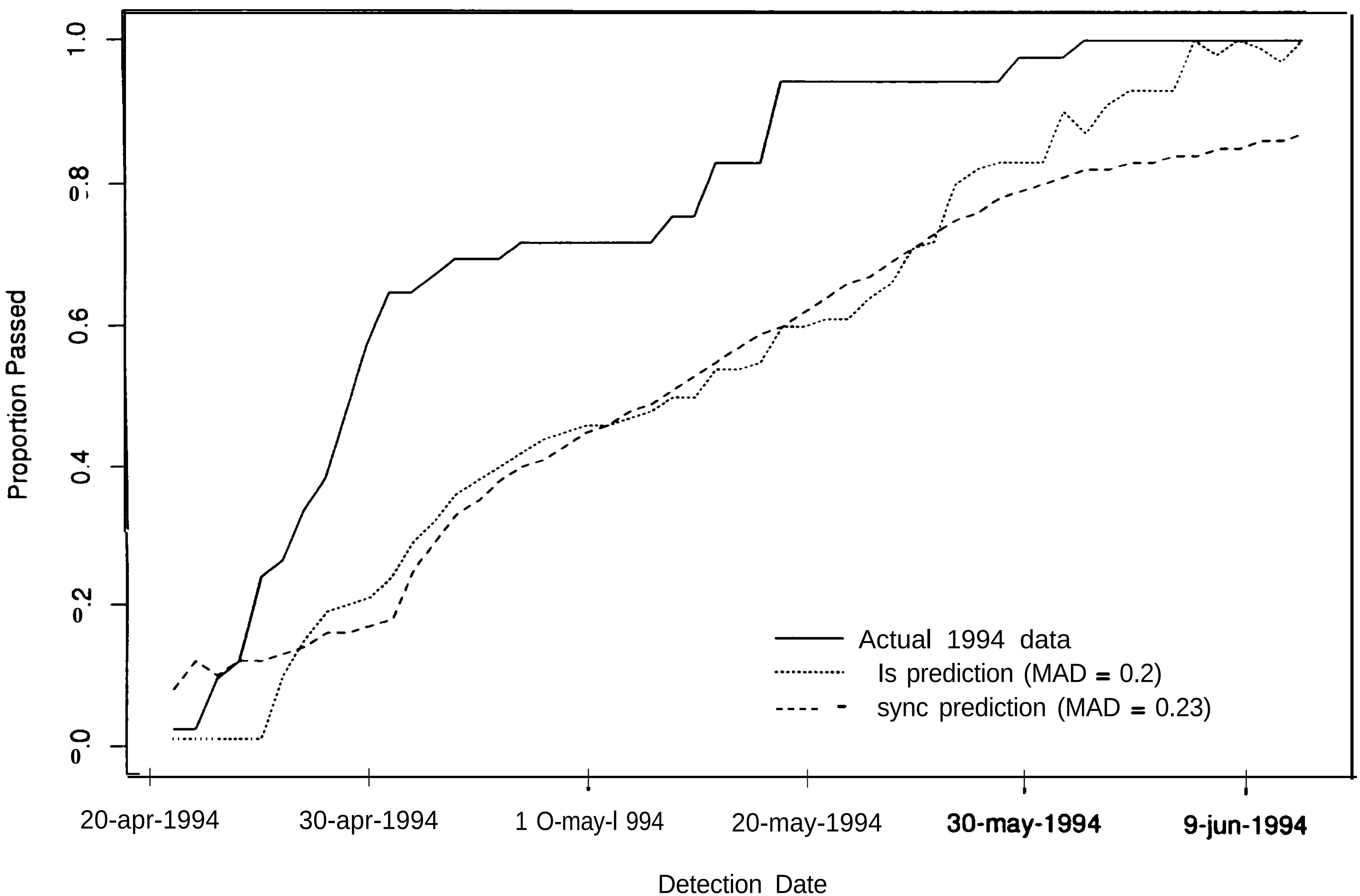


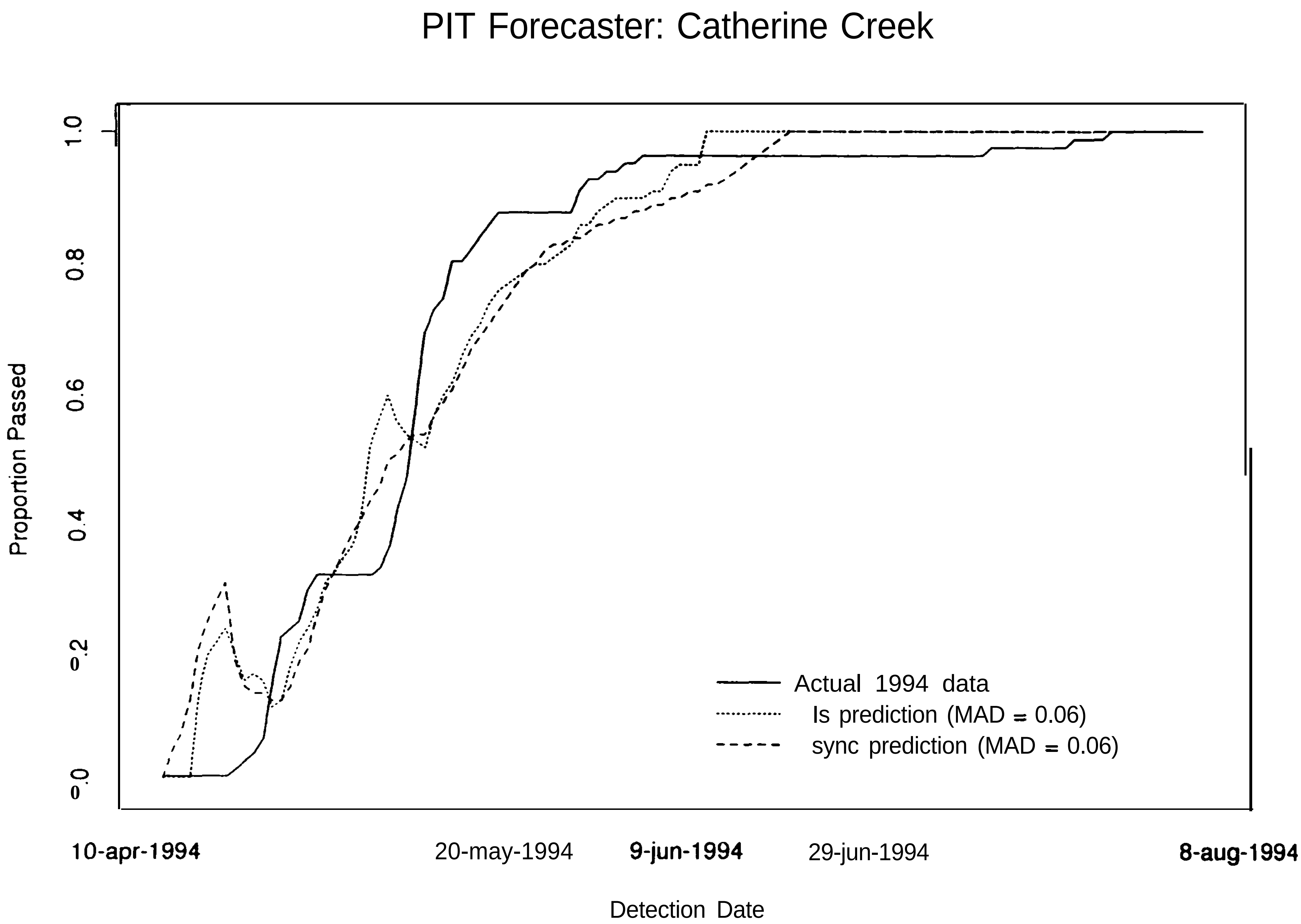
PIT Forecaster: Upper Middle Fork Salmon

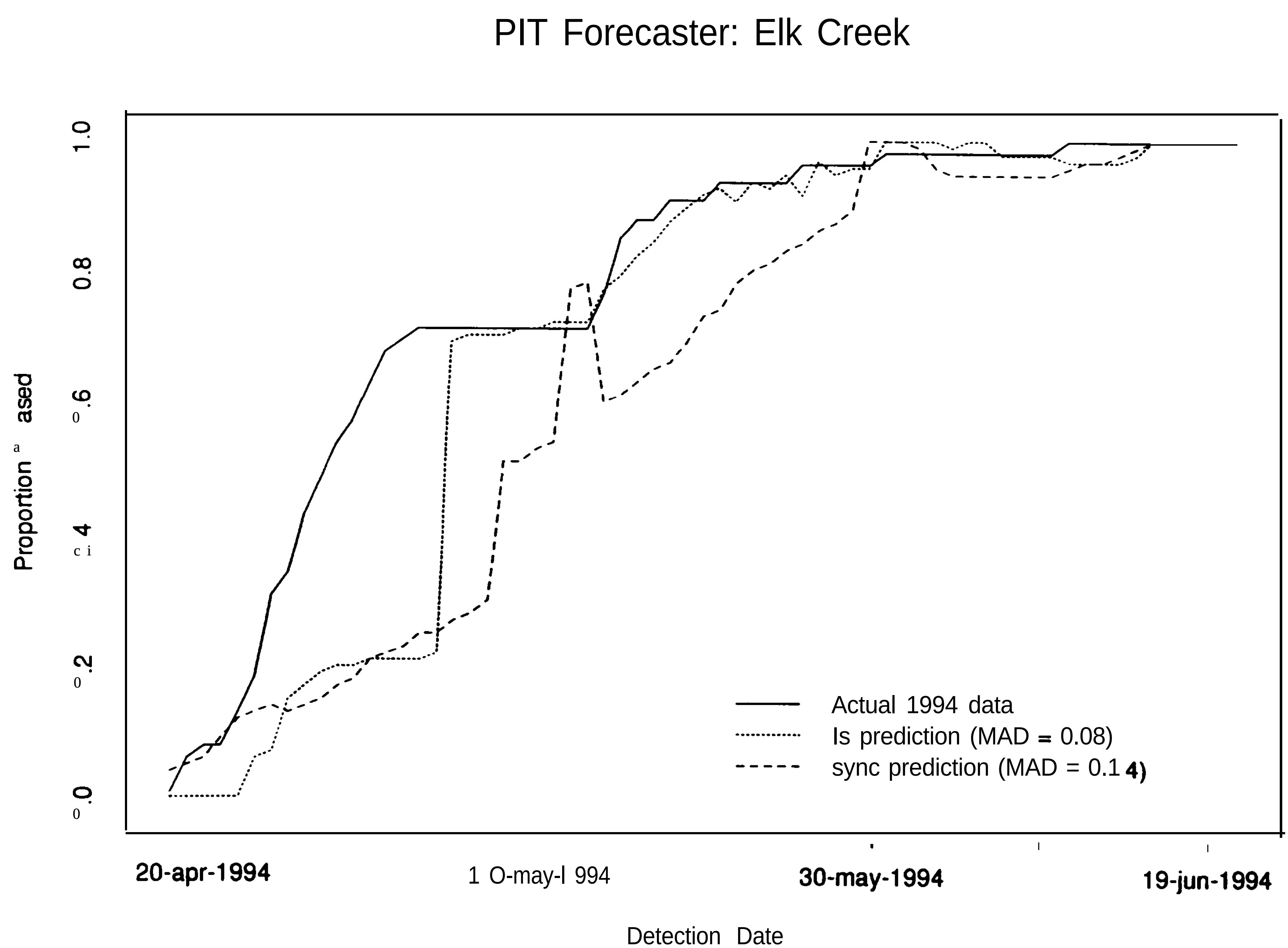



PIT Forecaster: Bear Valley Creek
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
20-apr-1994
1 O-may-l 994
30-may-1994
Detection Date
Actual 1994 data
ls prediction (MAD = 0.08)
sync prediction (MAD = 0.16)

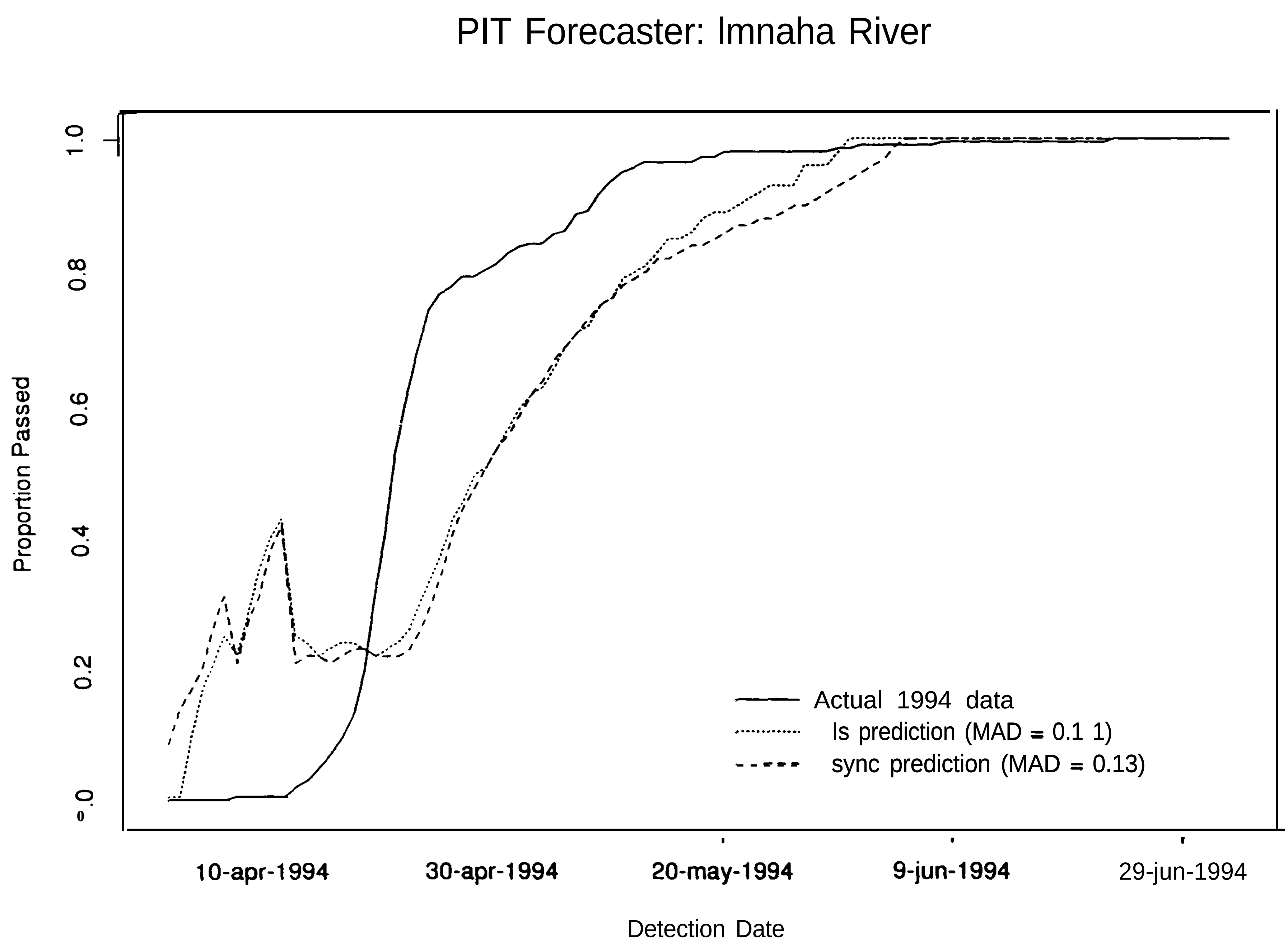
PIT Forecaster: Big Creek

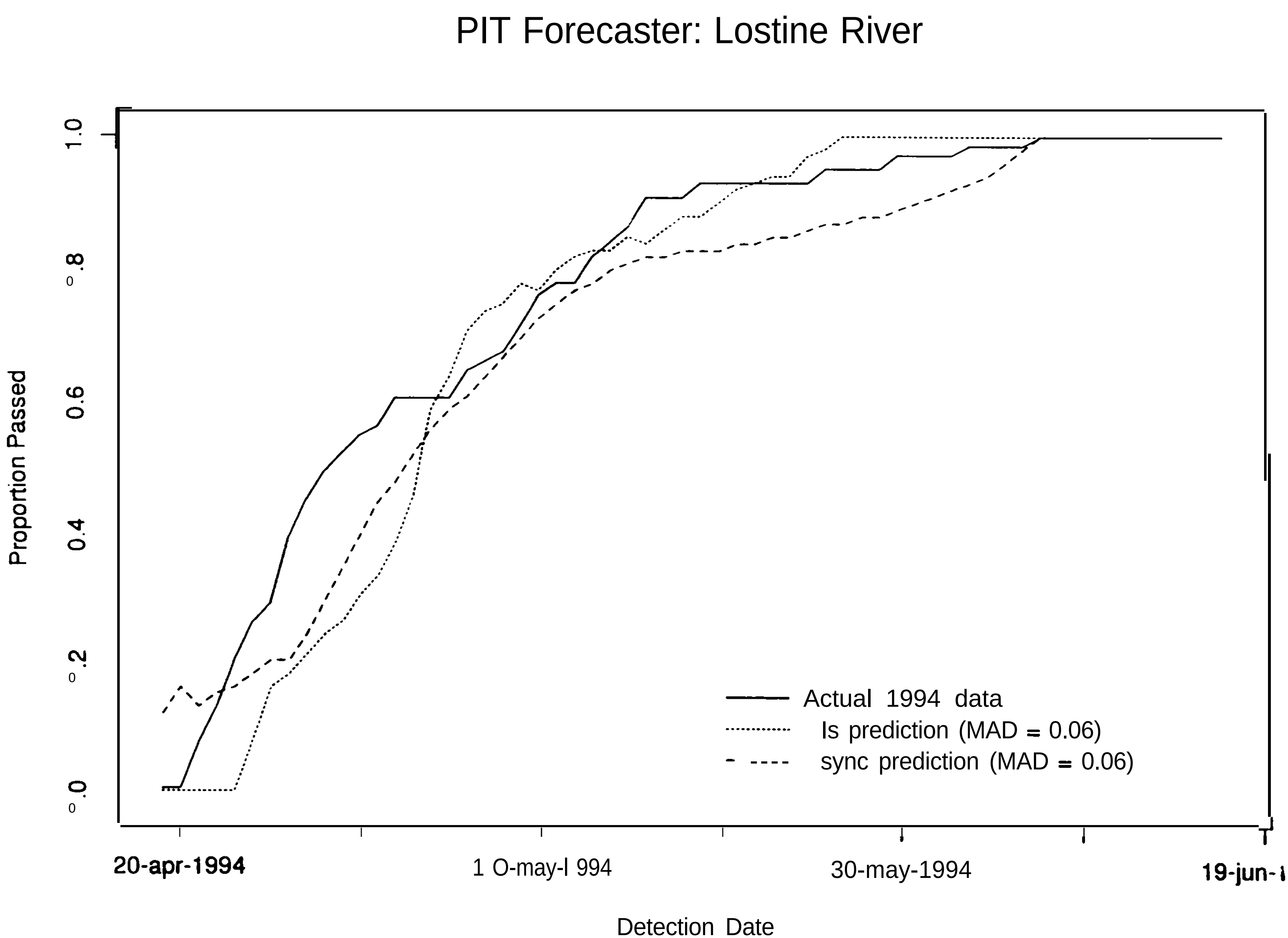



PIT Forecaster: Catherine Creek
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
Actual 1994 data
ls prediction (MAD = 0.06)
sync prediction (MAD = 0.06)
10-apr-1994
20-may-1994
9-jun-1994
29-jun-1994
8-aug-1994
Detection Date

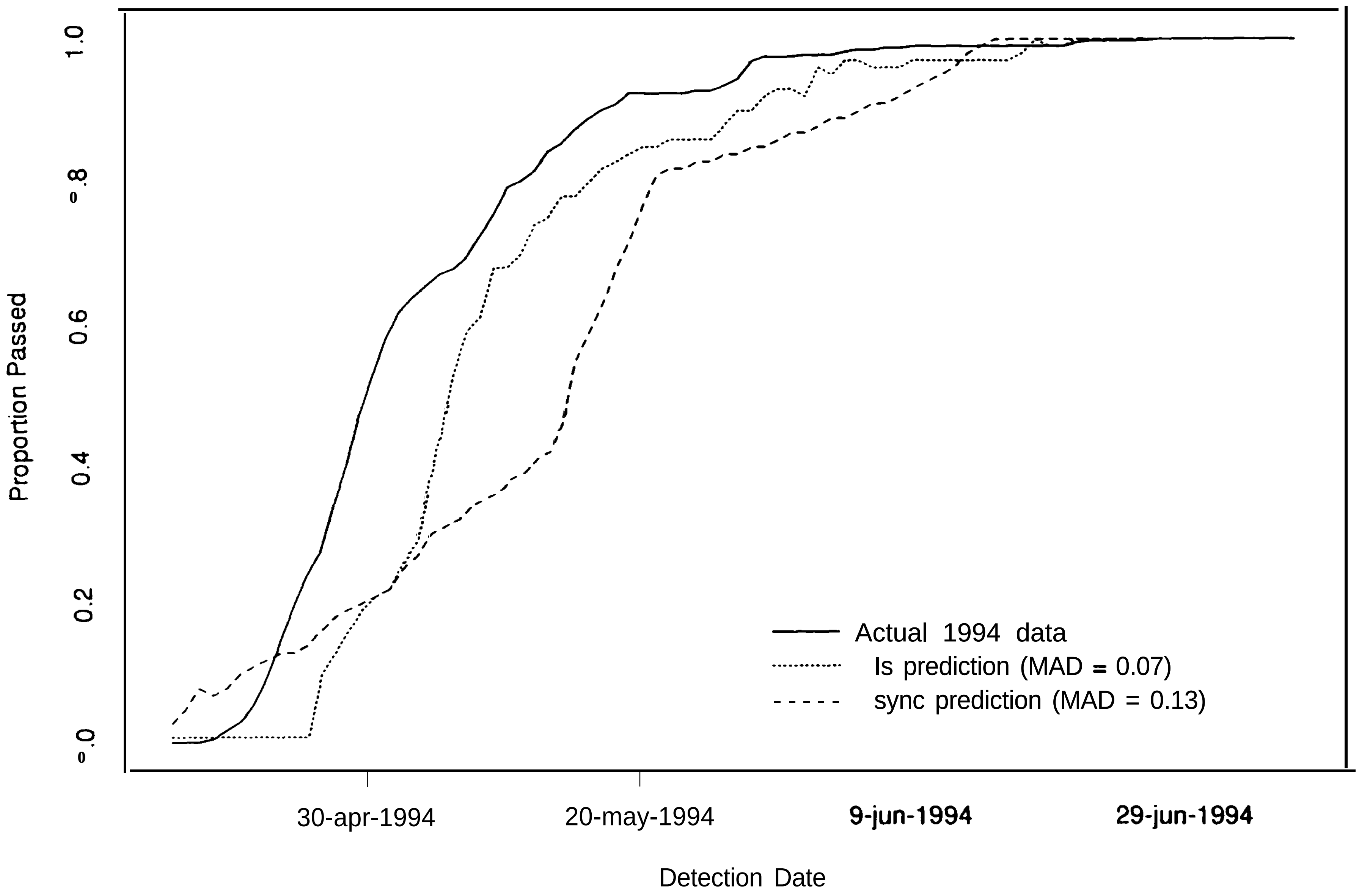


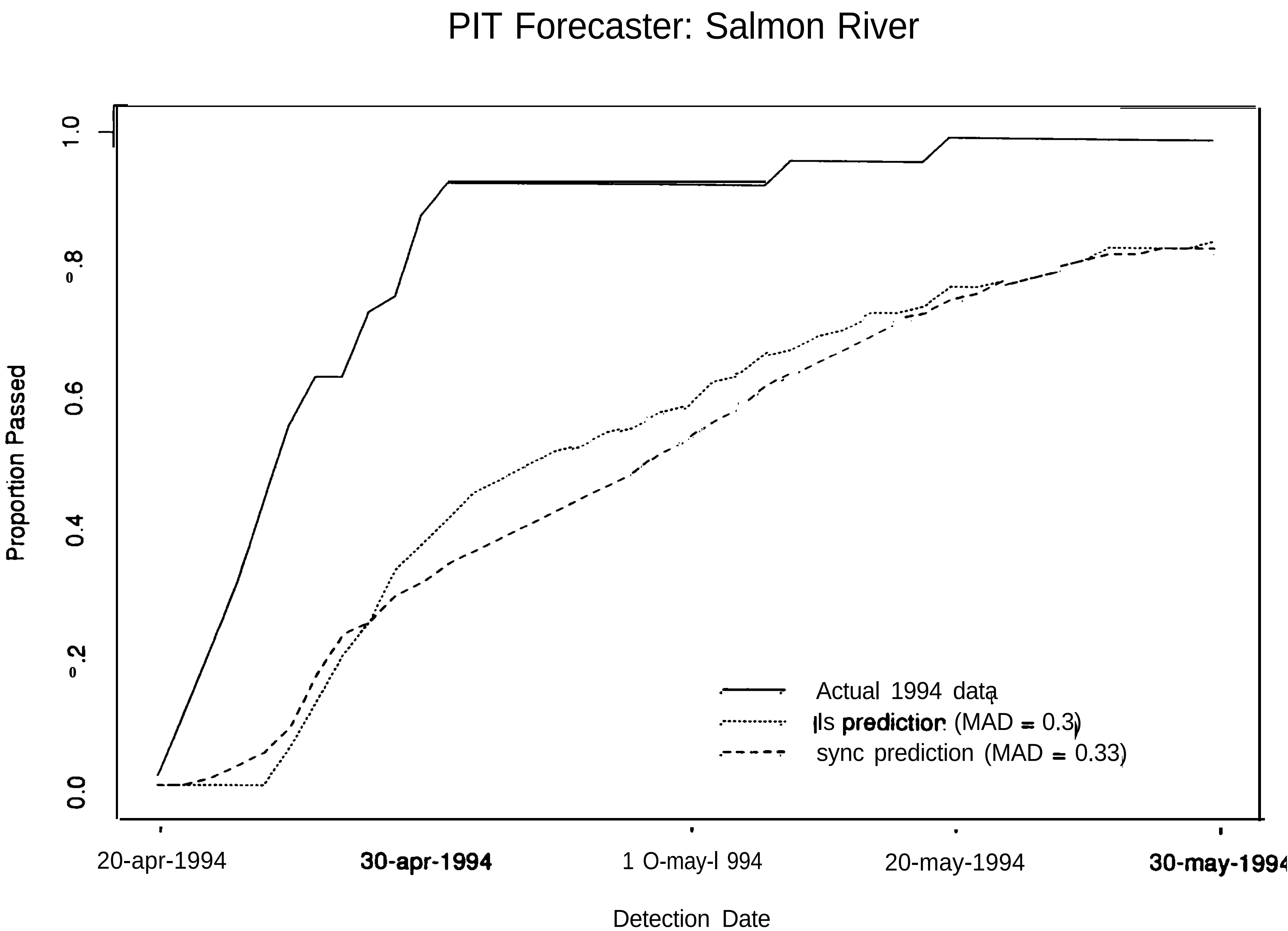
PIT Forecaster: Elk Creek
Proportion ased
0.0
0.2
c i 4
0.6
0.8
1.0
20-apr-1994
1 O-may-l 994
30-may-1994
19-jun-1994
Detection Date
Actual 1994 data
ls prediction (MAD = 0.08)
sync prediction (MAD = 0.1 4)

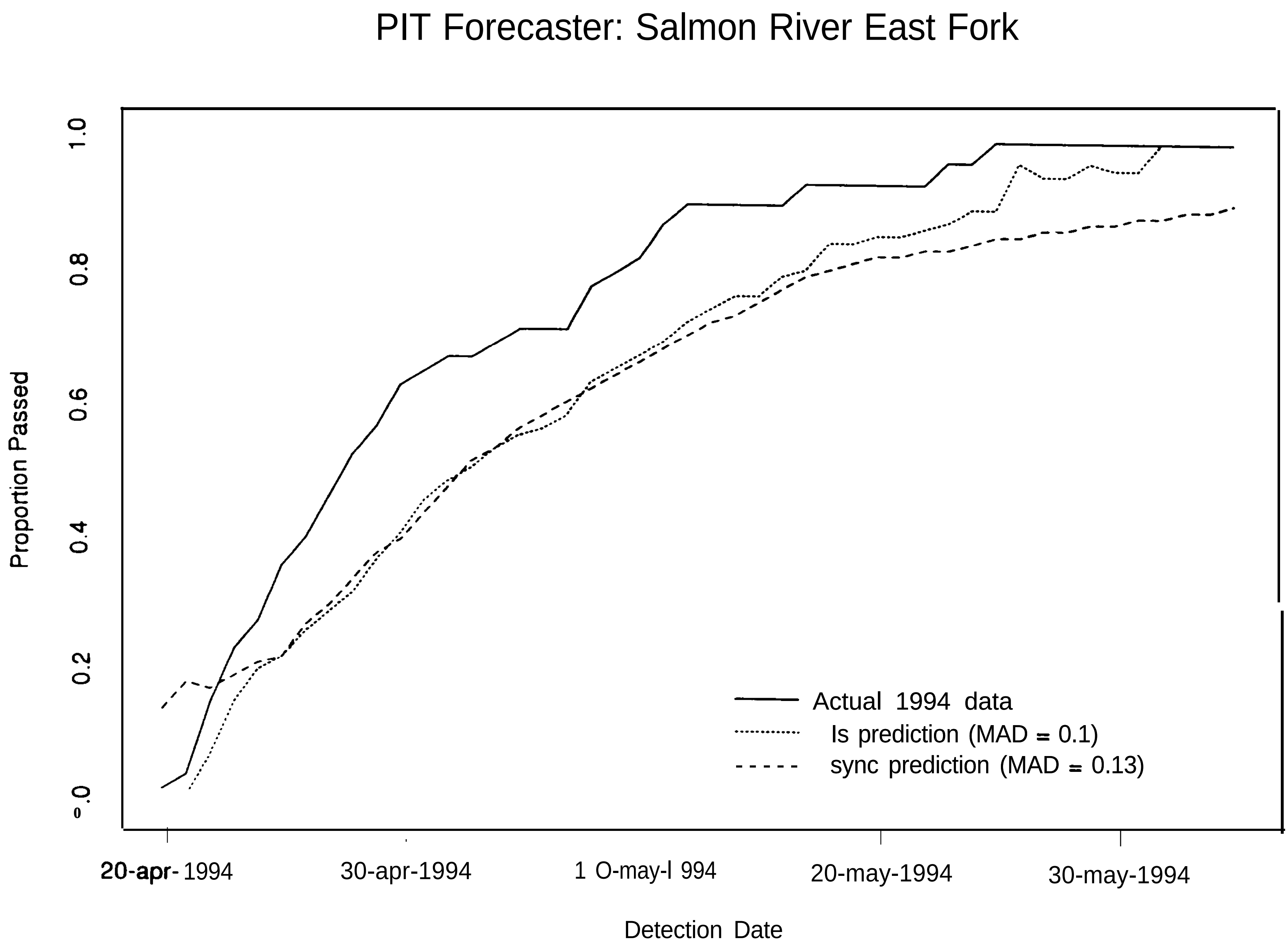

PIT Forecaster: Imnaha River
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
Actual 1994 data
ls prediction (MAD = 0.1 1)
sync prediction (MAD = 0.13)
10-apr-1994
30-apr-1994
20-may-1994
9-jun-1994
29-jun-1994
Detection Date

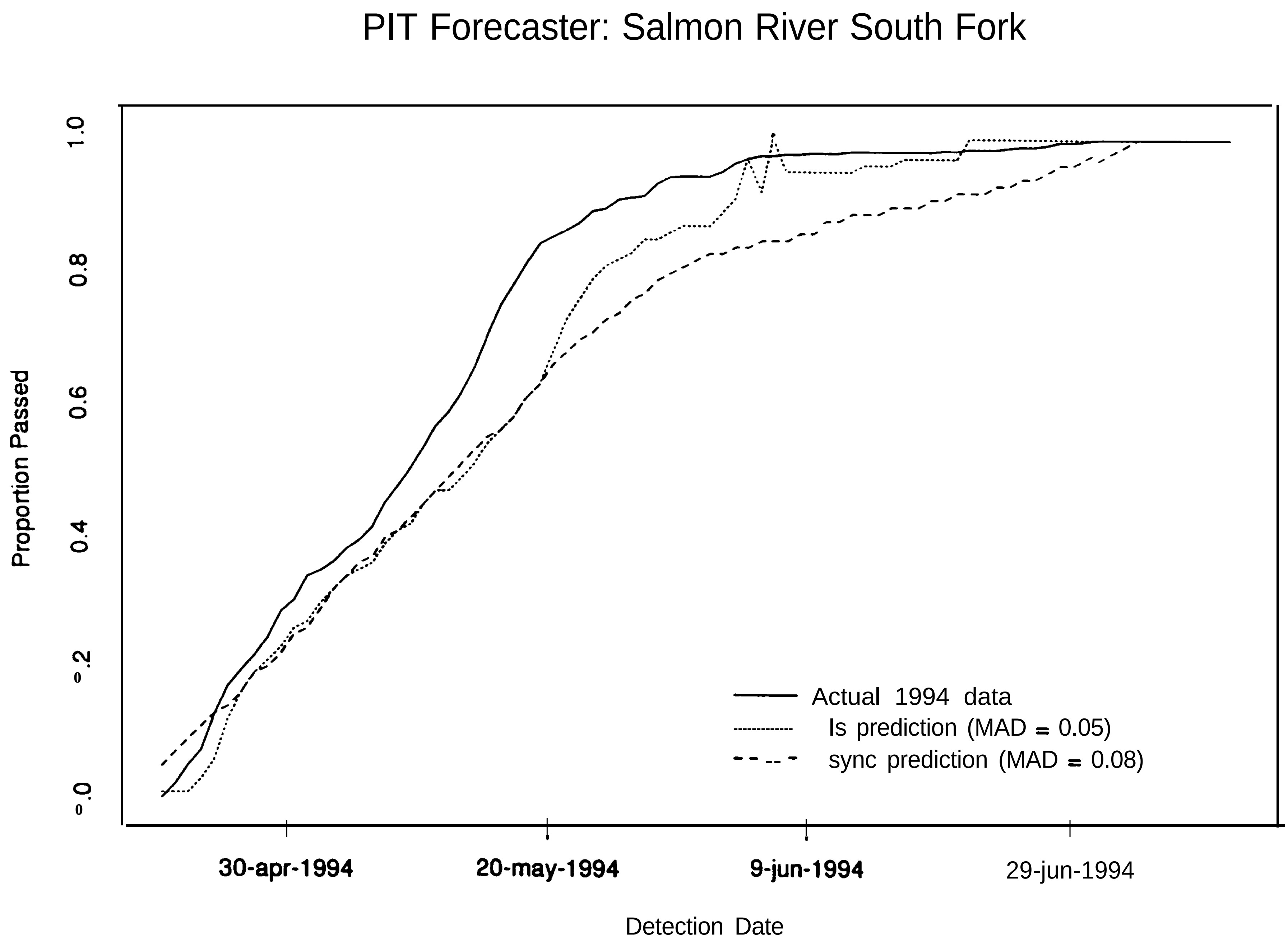

PIT Forecaster: Lostine River
Proportion Passed
1.0
0.8
0.6
0.4
0.2
0.0
20-apr-1994
1 O-may-l 994
30-may-1994
19-jun-
Detection Date
Actual 1994 data
ls prediction (MAD = 0.06)
sync prediction (MAD = 0.06)

PIT Forecaster: Marsh Creek

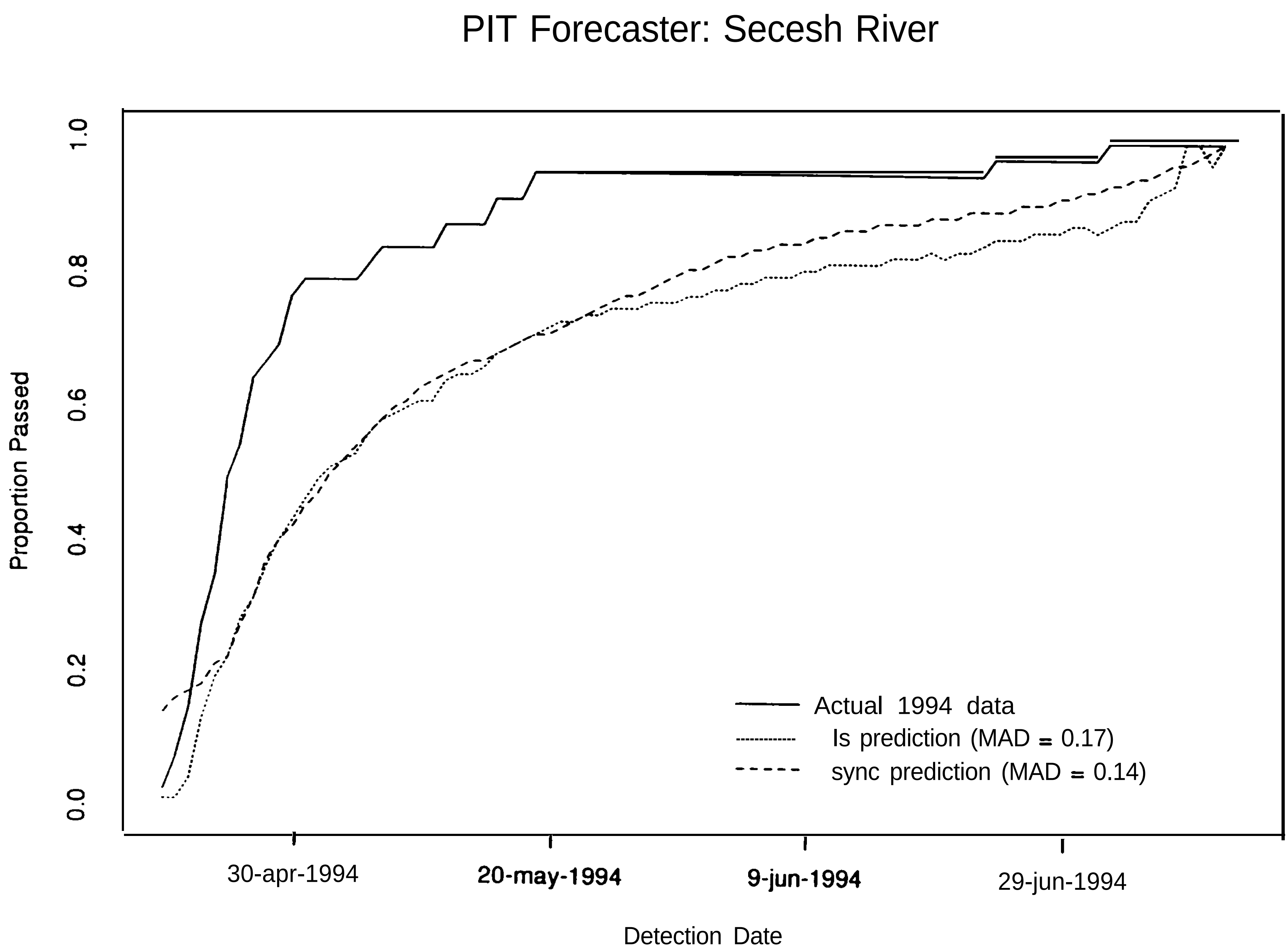


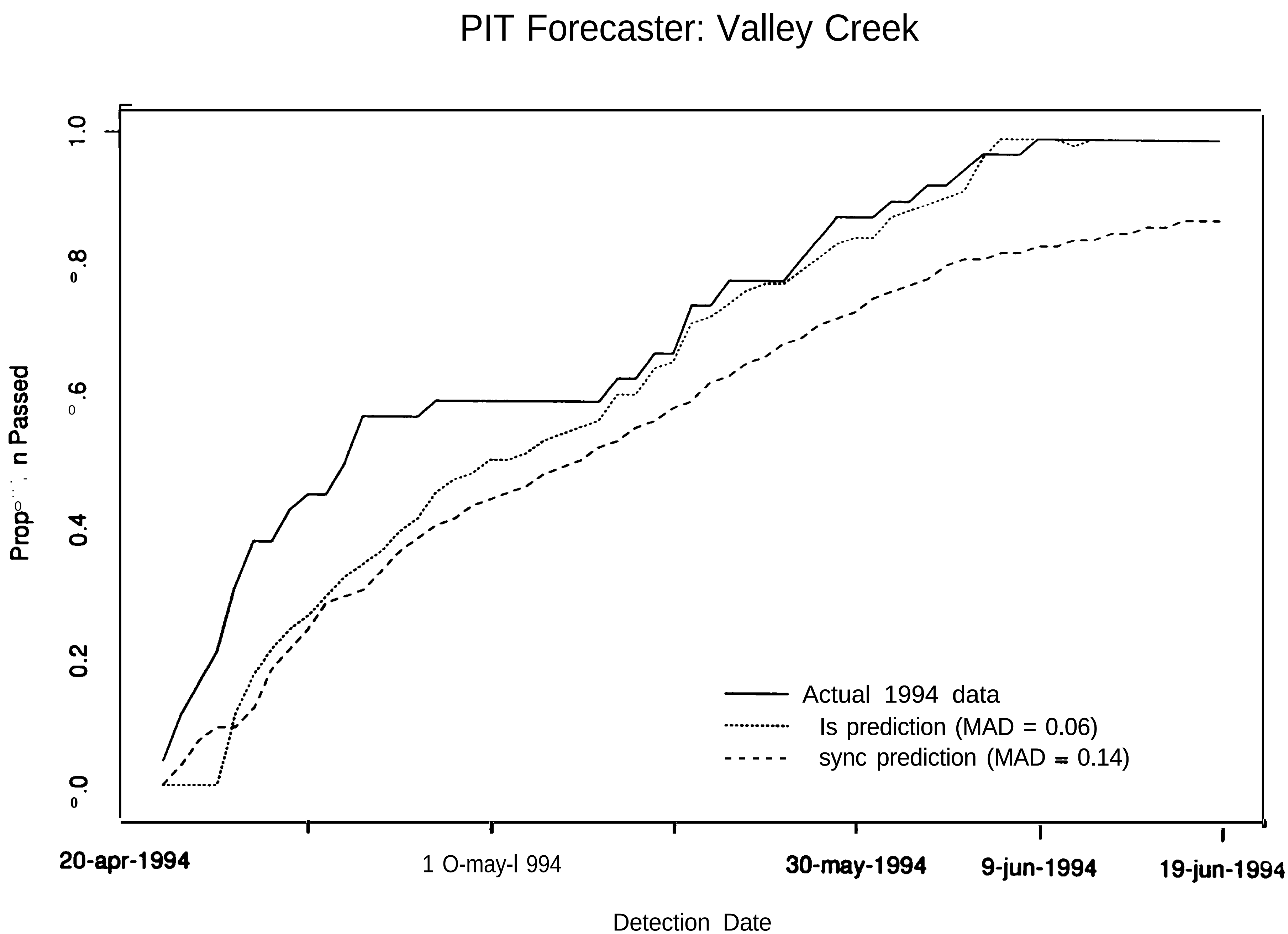

PIT Forecaster: Salmon River
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
Actual 1994 data
lls prediction (MAD = 0.3)
sync prediction (MAD = 0.33)
20-apr-1994
30-apr-1994
10-may-1994
20-may-1994
30-may-1994
Detection Date


PIT Forecaster: Salmon River East Fork
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
20-apr-1994
30-apr-1994
1 O-may-l 994
20-may-1994
30-may-1994
Detection Date
Actual 1994 data
ls prediction (MAD = 0.1)
sync prediction (MAD = 0.13)



PIT Forecaster: Salmon River South Fork
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
Actual 1994 data
ls prediction (MAD = 0.05)
sync prediction (MAD = 0.08)
30-apr-1994
20-may-1994
9-jun-1994
29-jun-1994
Detection Date

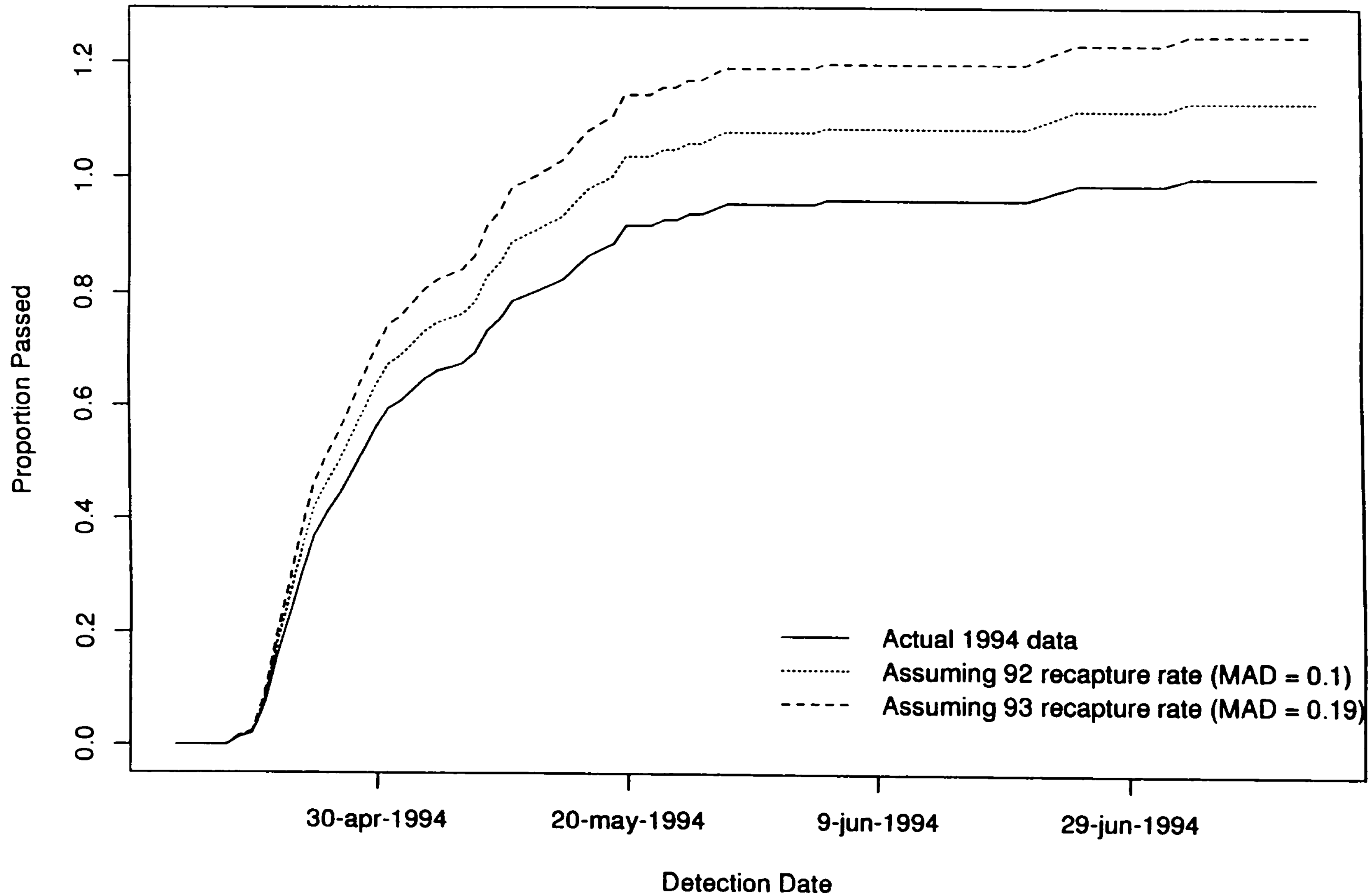

PIT Forecaster: Secesh River
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
30-apr-1994
20-may-1994
9-jun-1994
29-jun-1994
Detection Date
Actual 1994 data
ls prediction (MAD = 0.17)
sync prediction (MAD = 0.14)



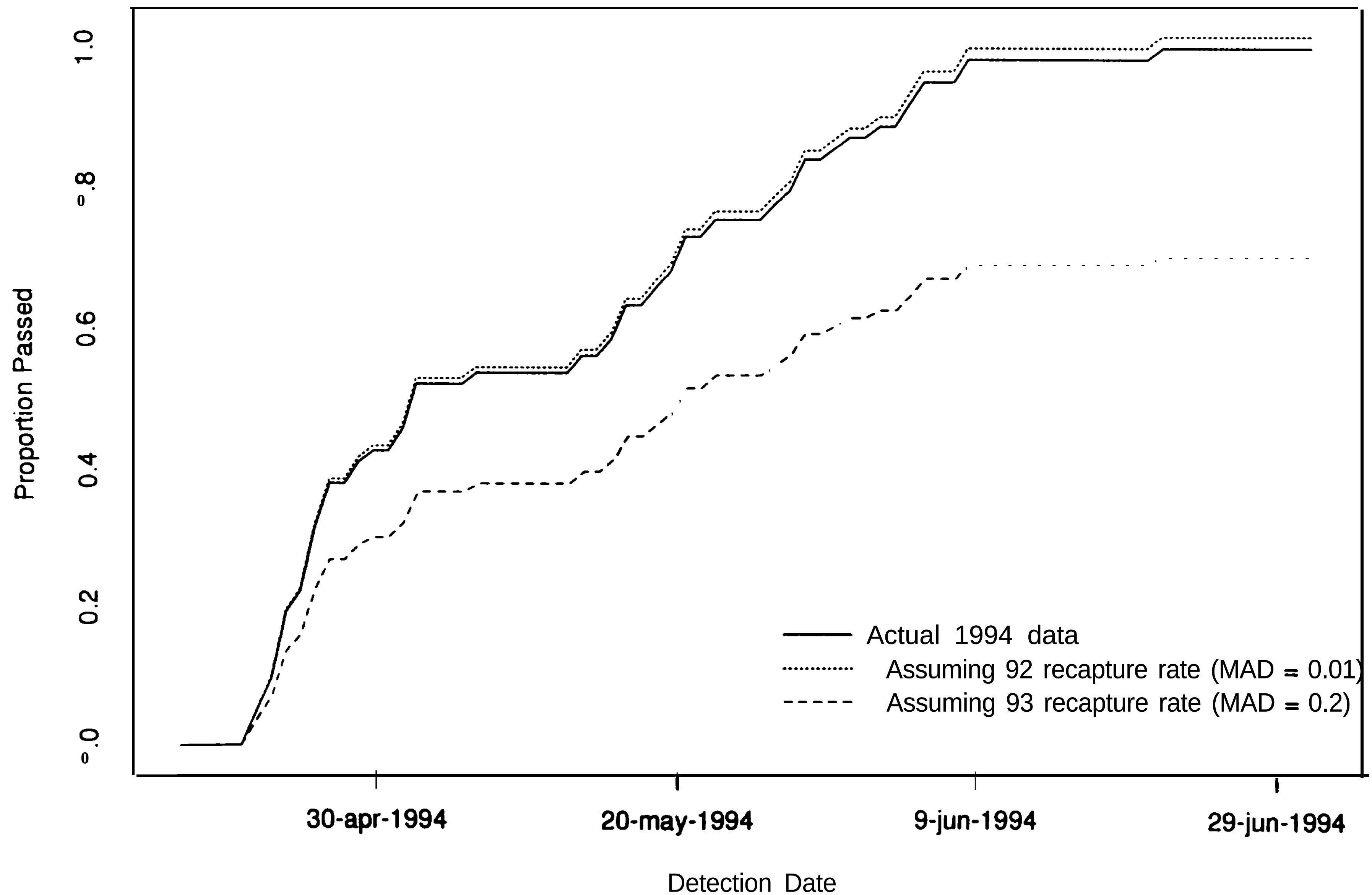
PIT Forecaster: Valley Creek
Prop n Passed
0.0
0.2
0.4
0.6
0.8
1.0
20-apr-1994
10-may-1994
30-may-1994
9-jun-1994
19-jun-1994
Detection Date
Actual 1994 data
ls prediction (MAD = 0.06)
sync prediction (MAD = 0.14)

Alternative Method #1 Plots

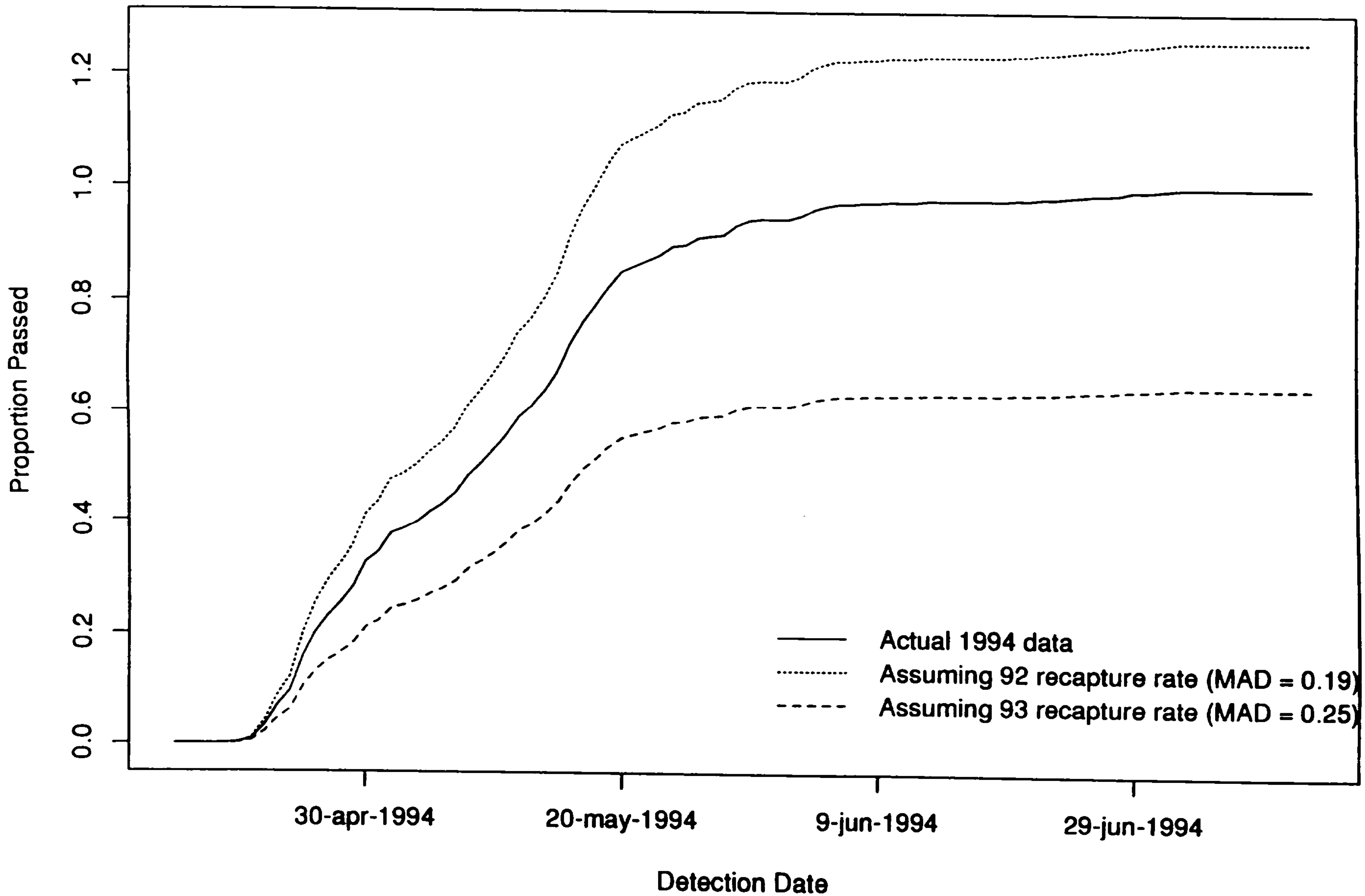
Alternative Method #1: Pahsimeroi



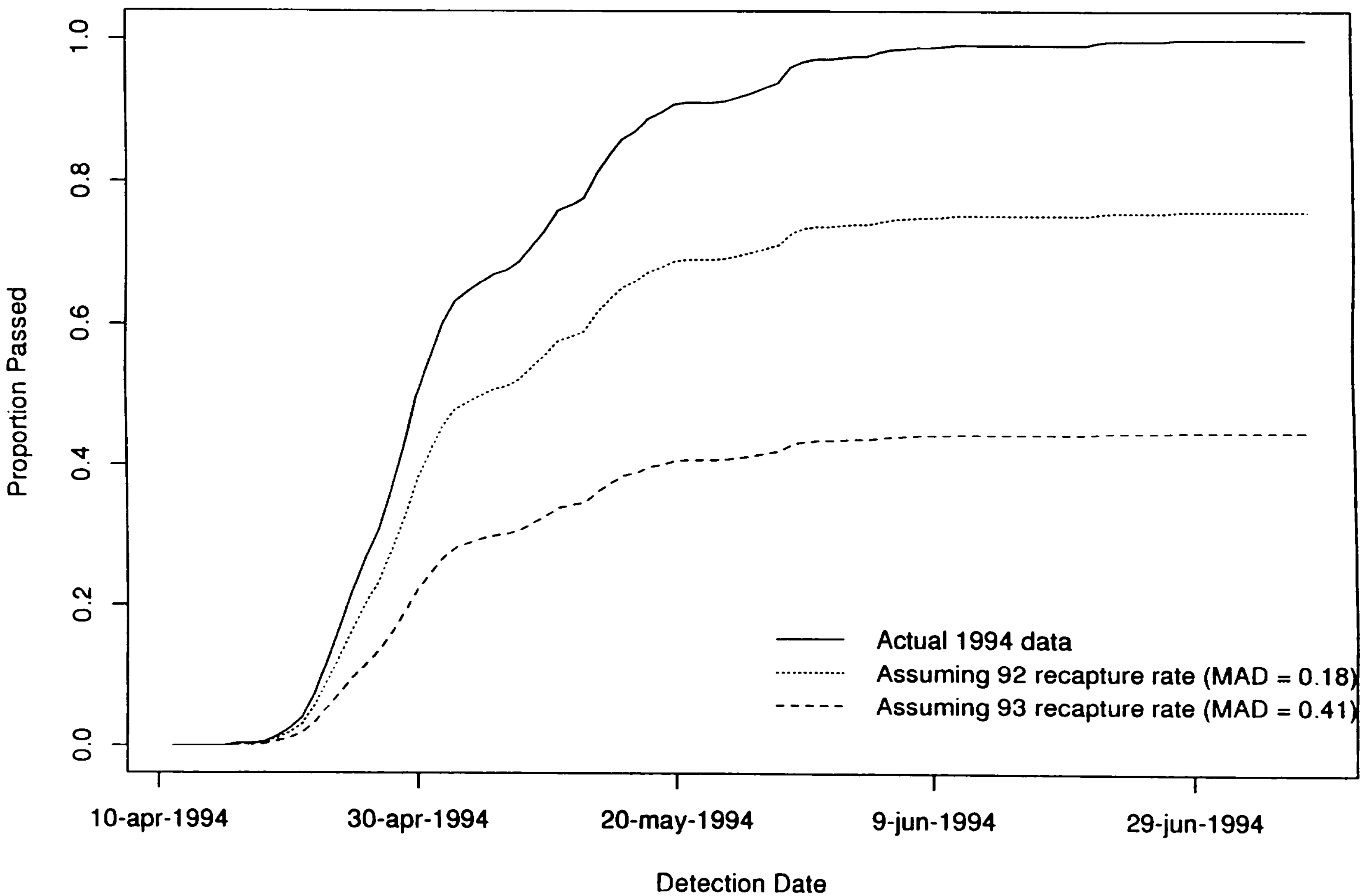
Alternative Method #1: Upper Salmon



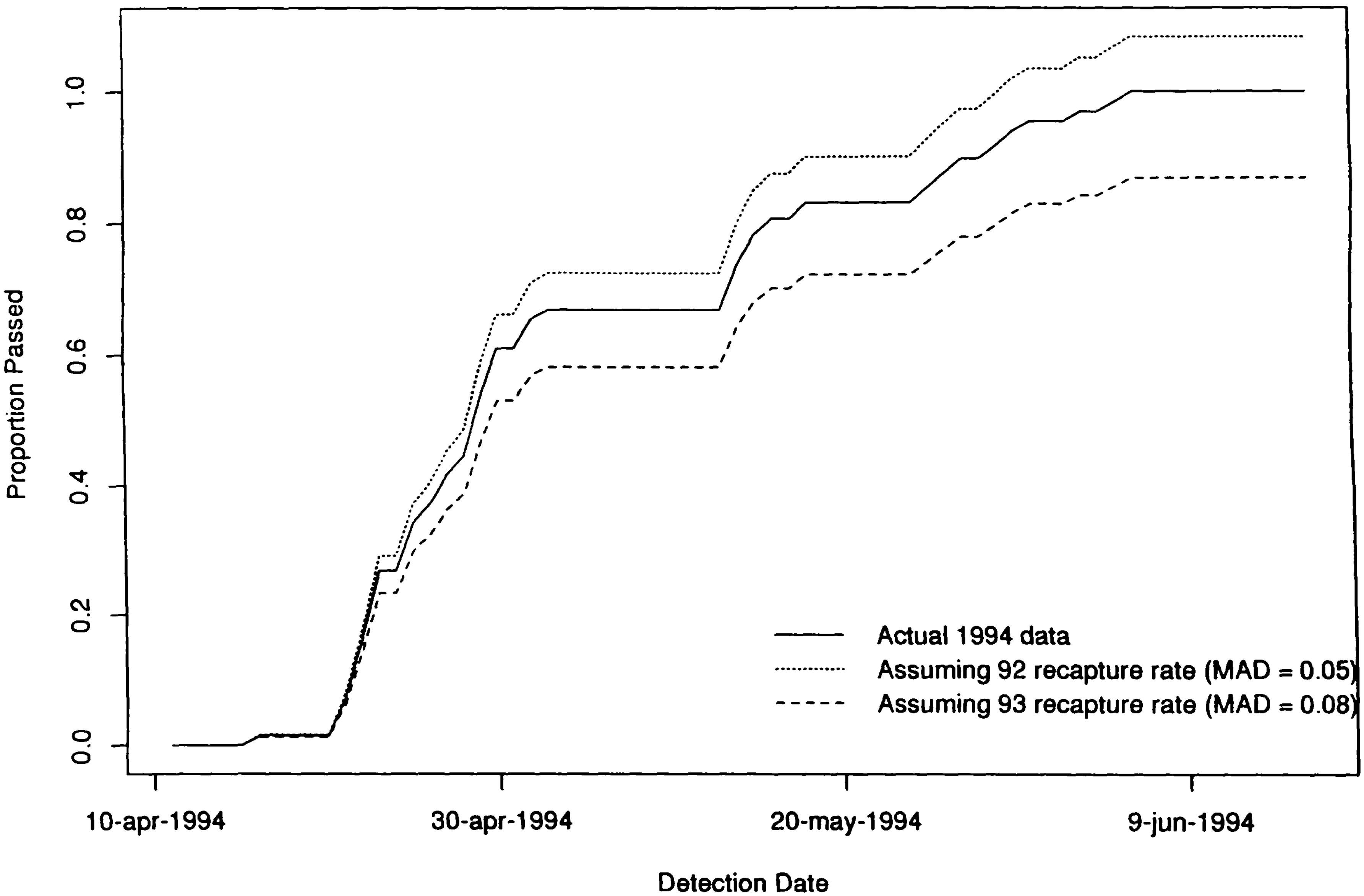
Alternative Method #1: South Fork Salmon



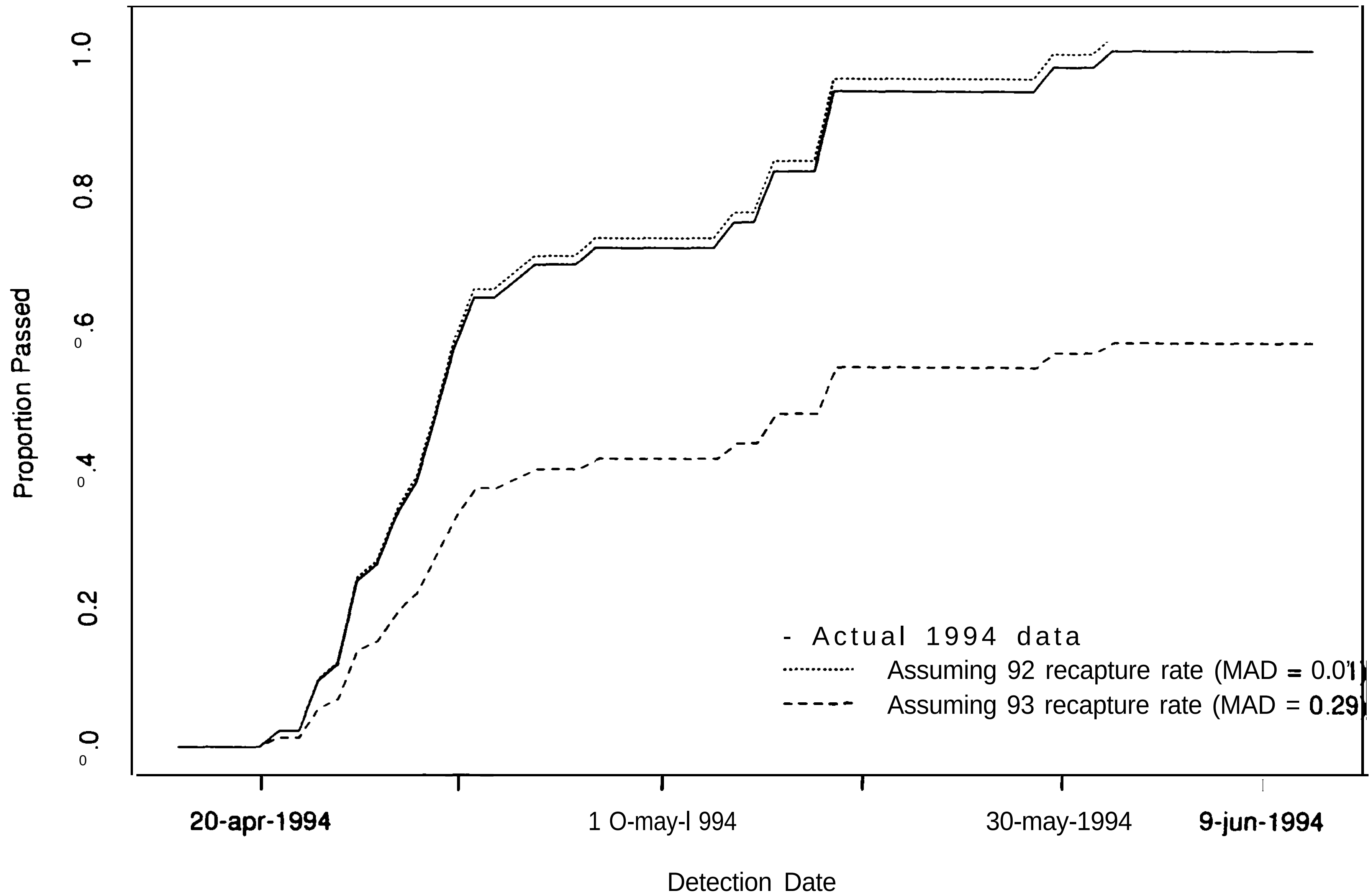
Alternative Method #1: Upper Middle Fork Salmon



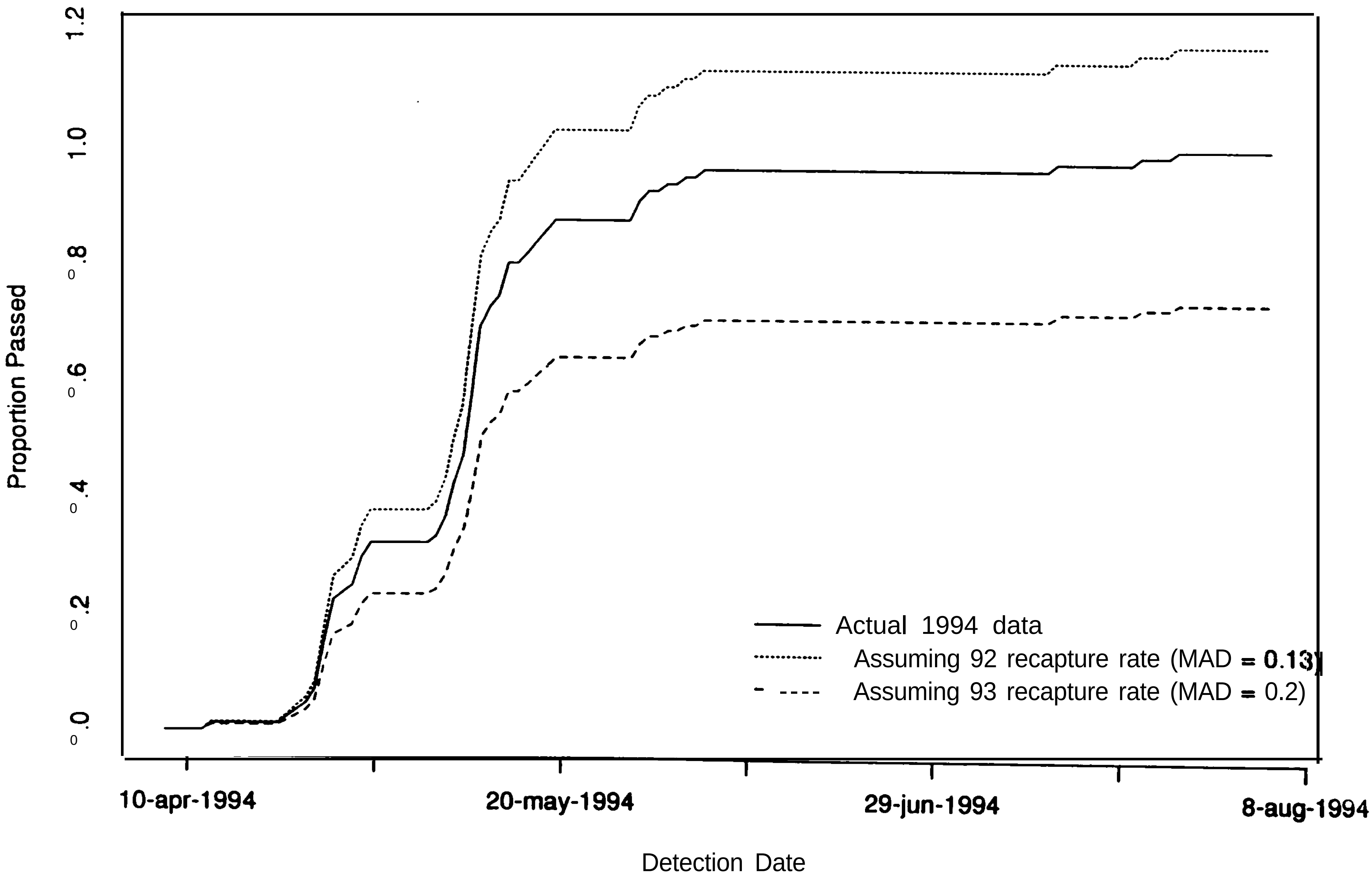
Alternative Method #1: Bear Valley Creek



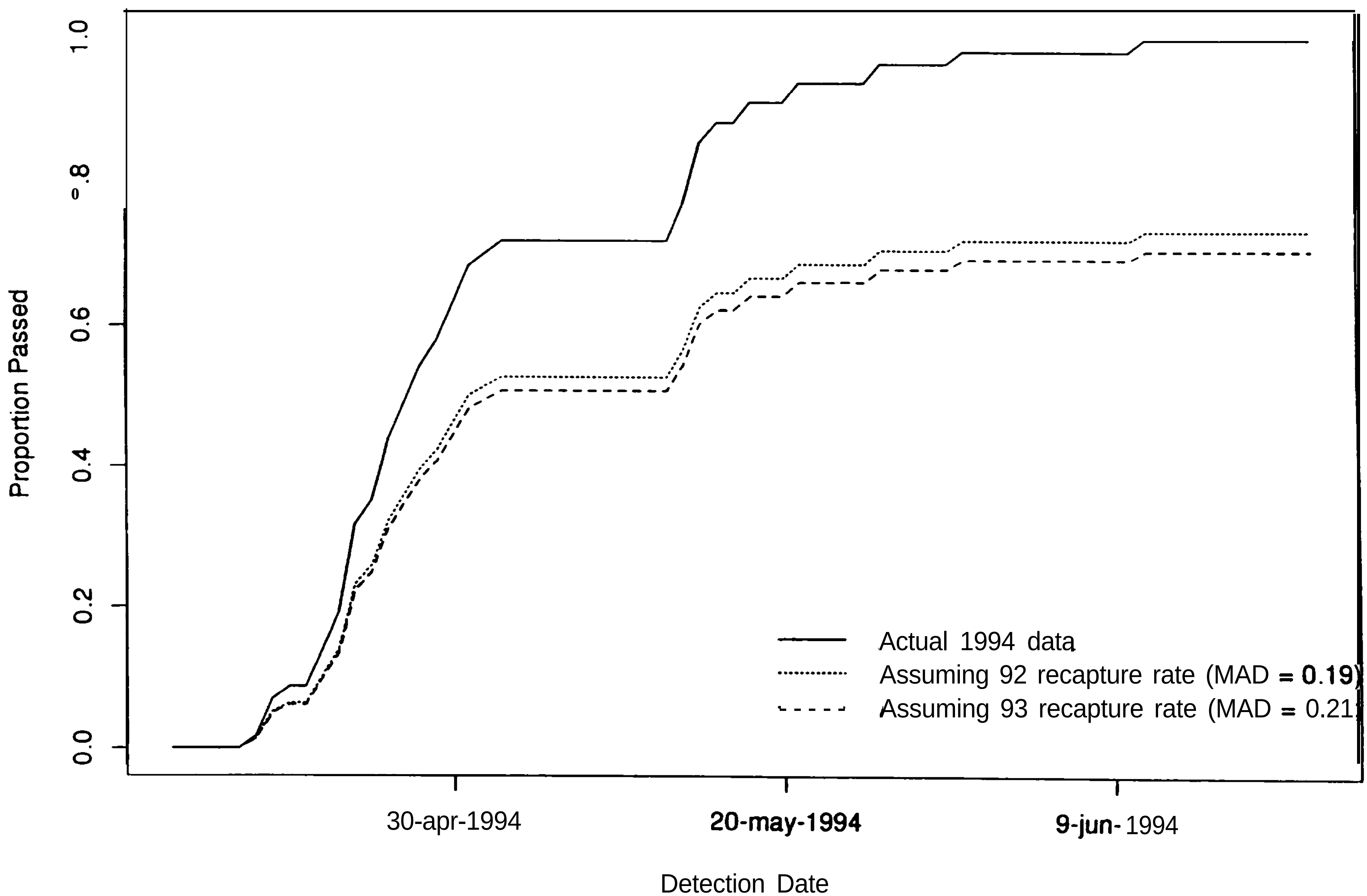
Alternative Method #1 : Big Creek



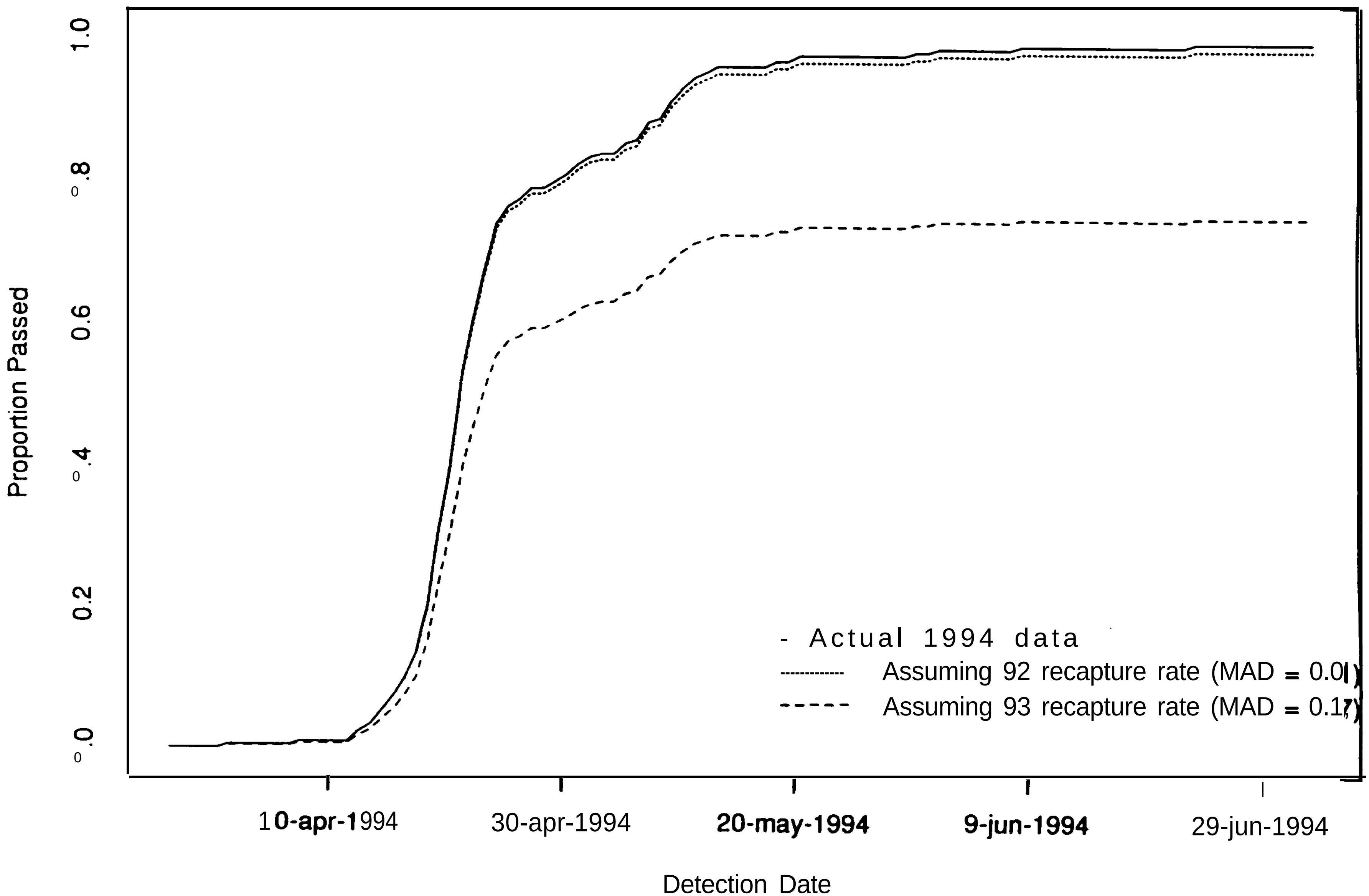
Alternative Method #1: Catherine Creek



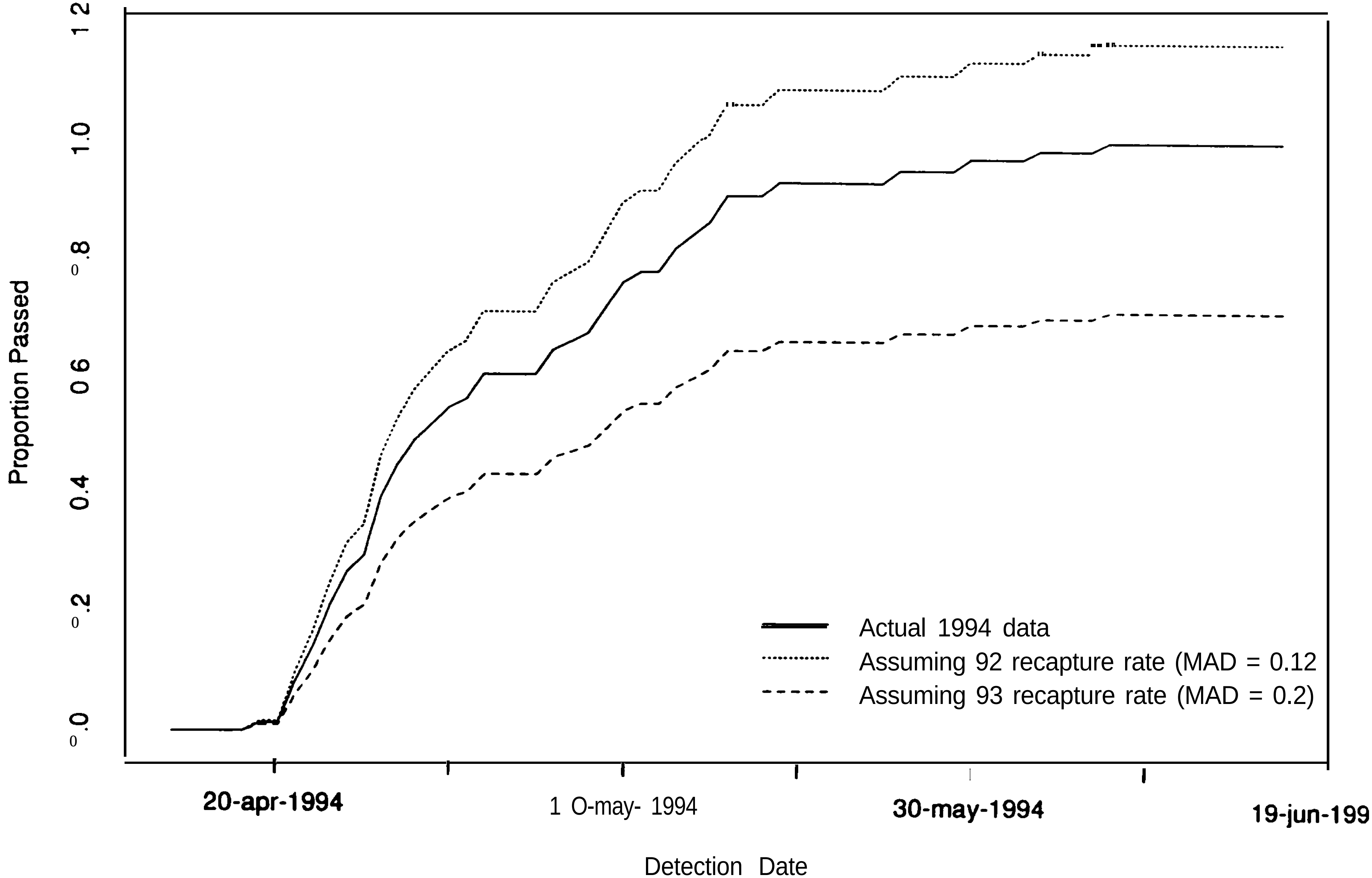
Alternative Method #1: Elk Creek



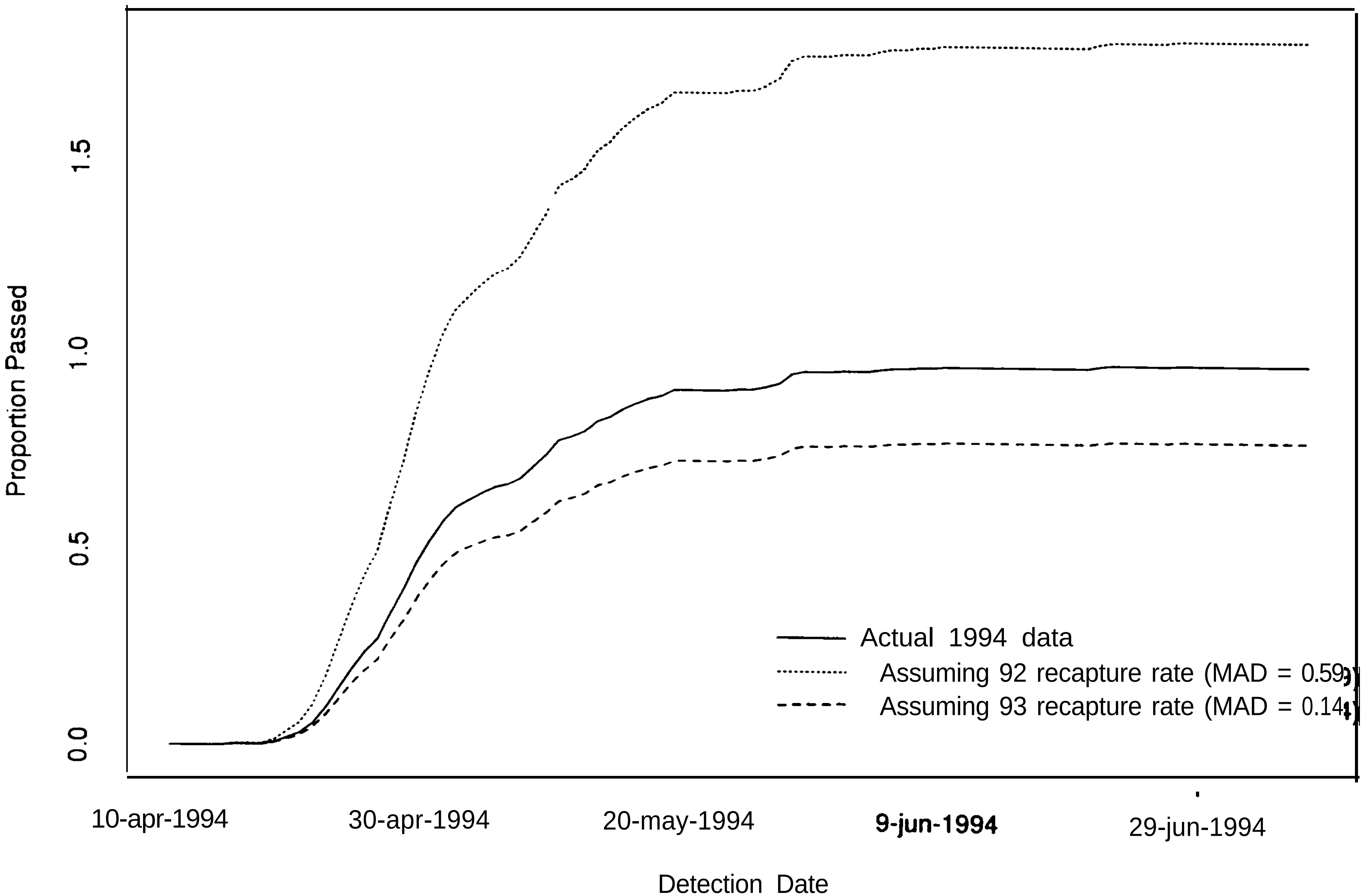
Alternative Method #1 : Imnaha River

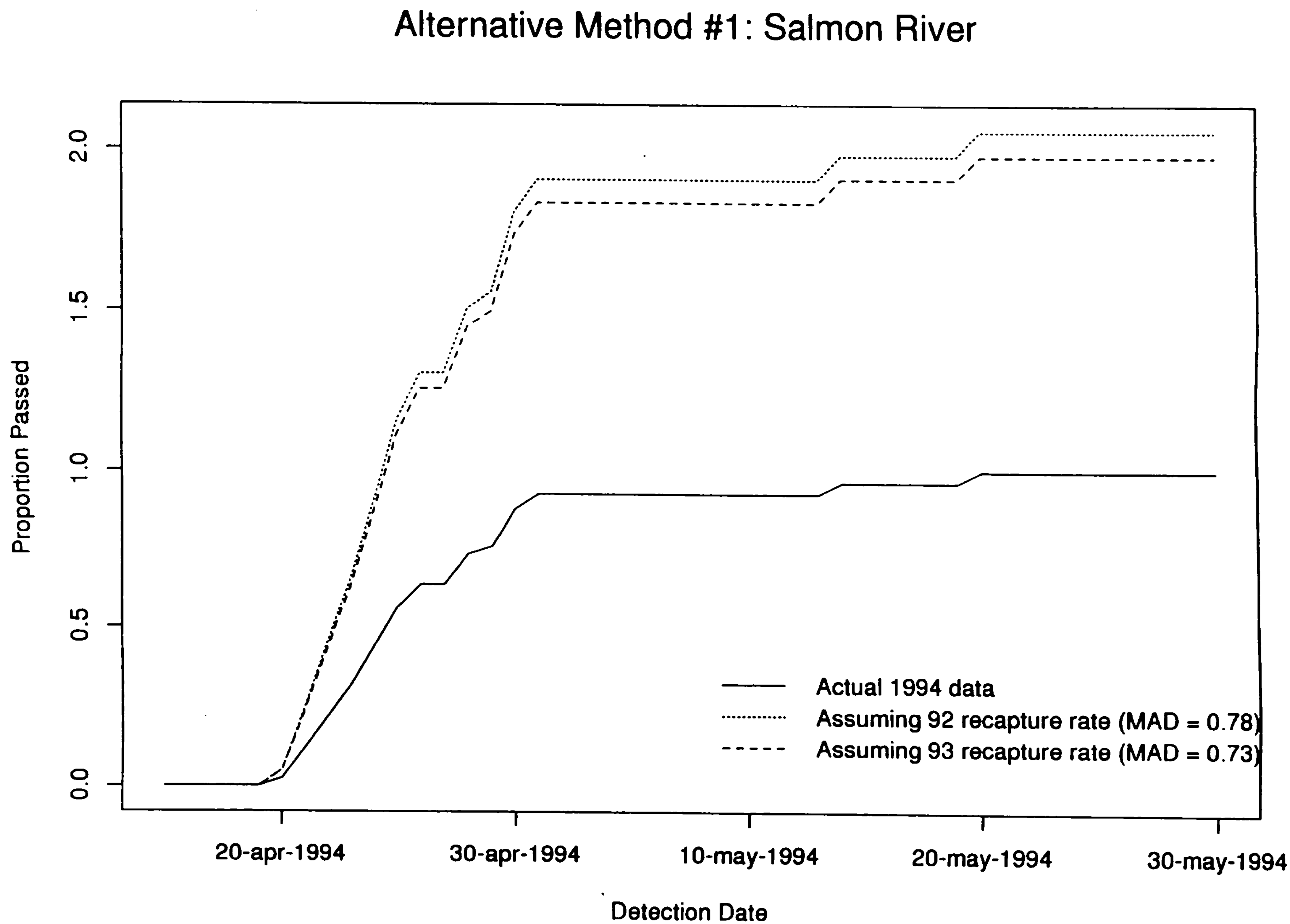


Alternative Method #1 : Lostine River

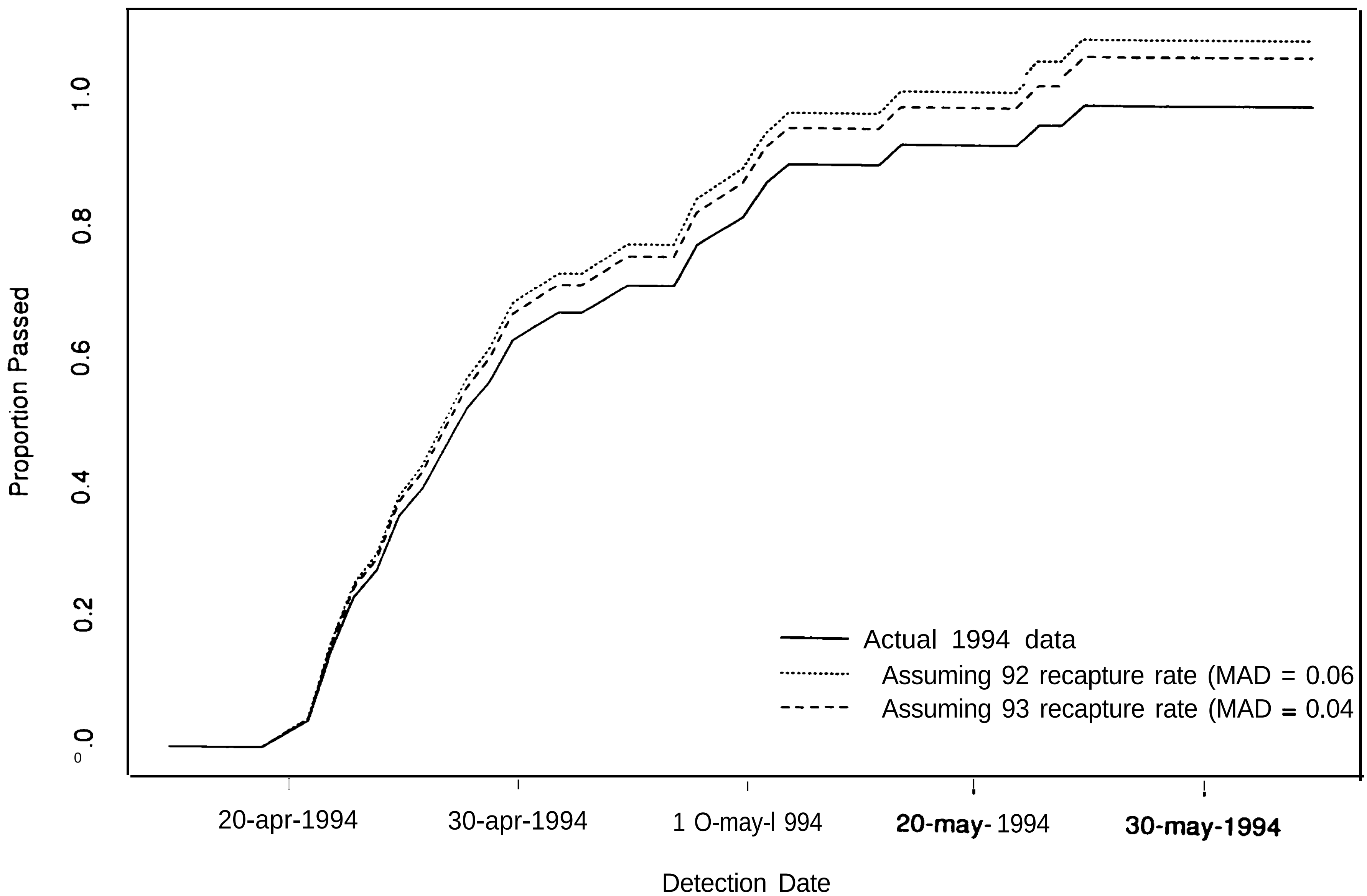


Alternative Method #1: Marsh Creek

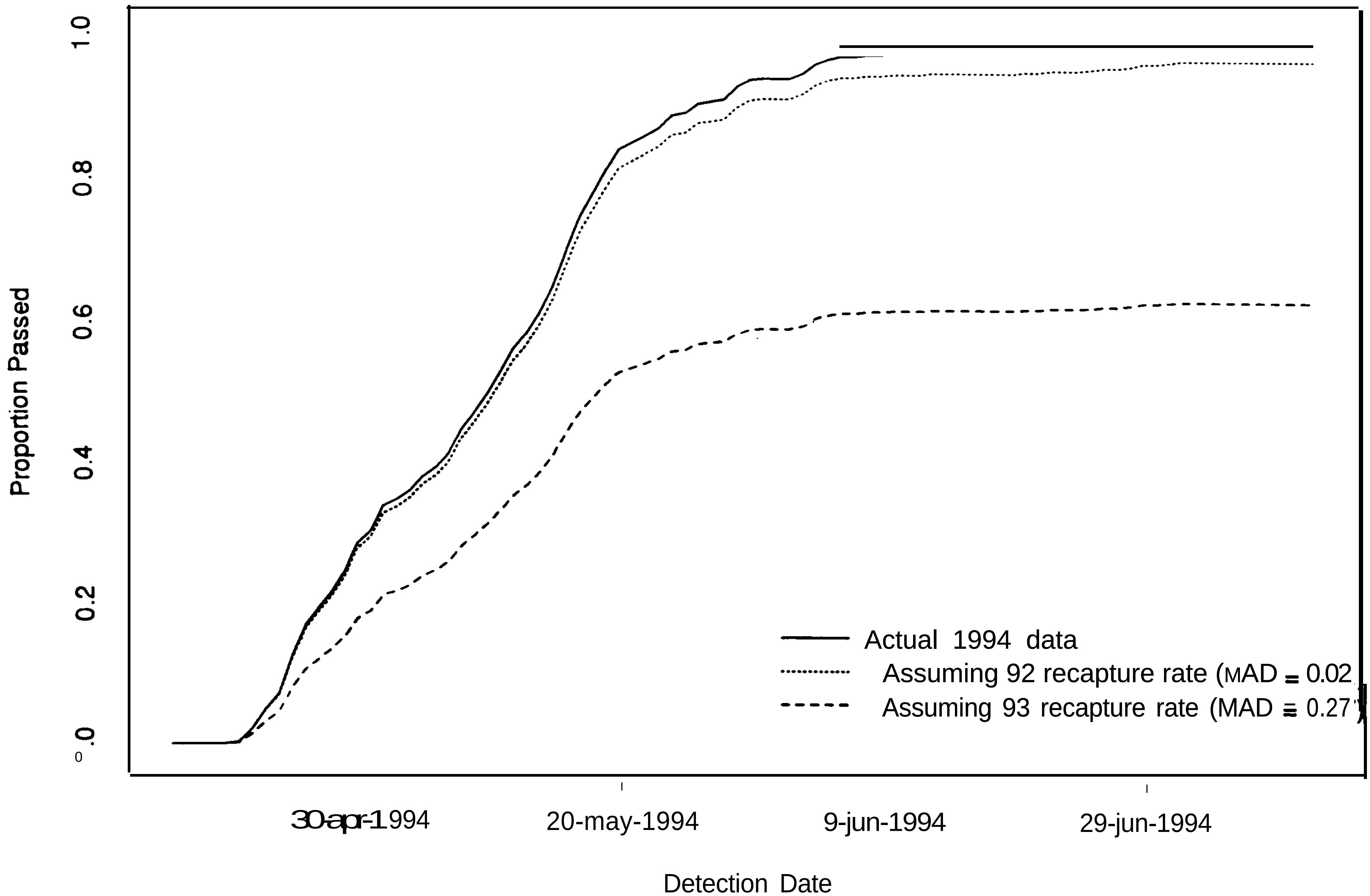


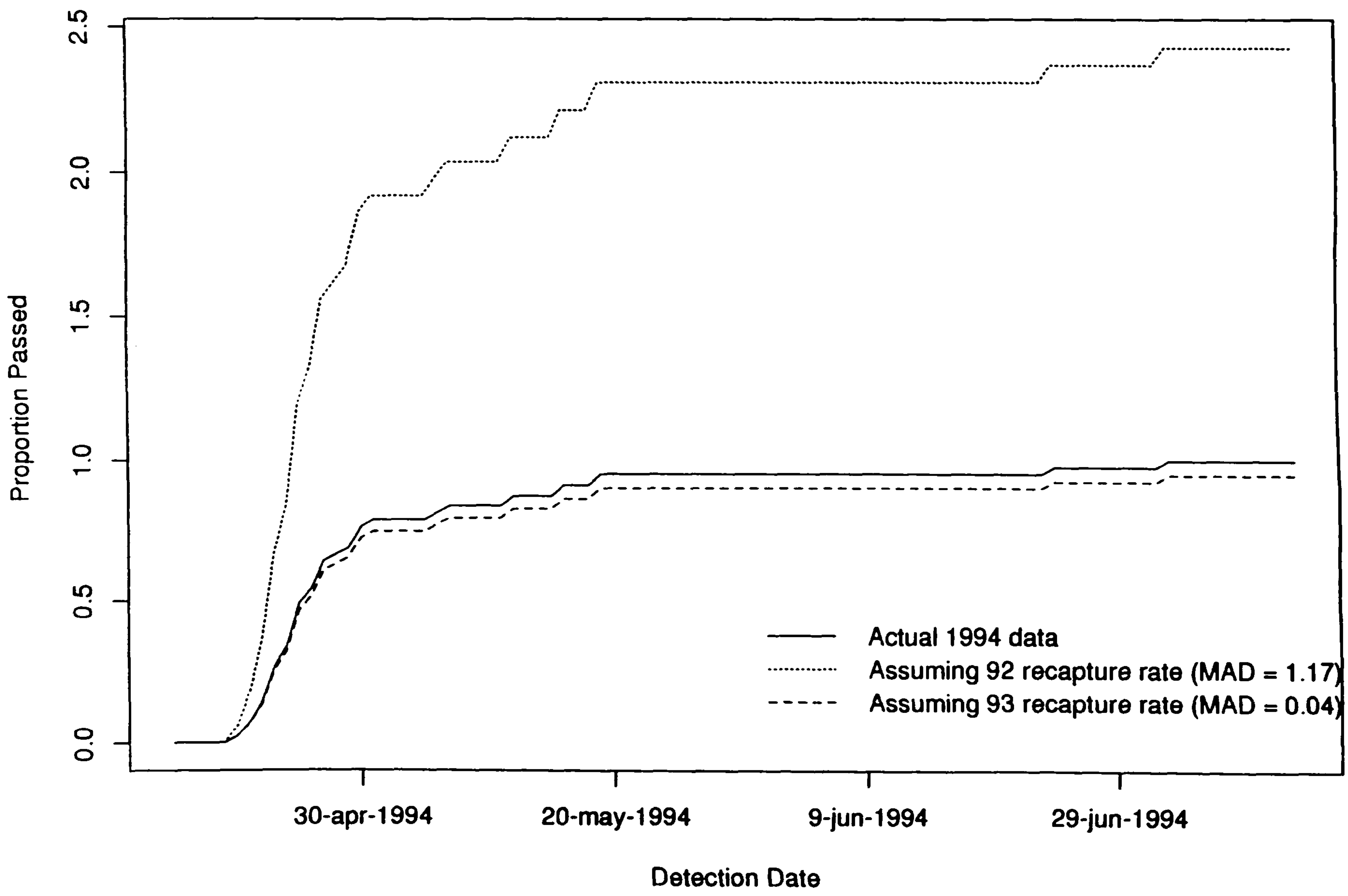

Alternative Method #1: Salmon River
Proportion Passed
0.0
0.5
1.0
1.5
2.0
Detection Date
20-apr-1994
30-apr-1994
10-may-1994
20-may-1994
30-may-1994
Actual 1994 data
Assuming 92 recapture rate (MAD = 0.78)
Assuming 93 recapture rate (MAD = 0.73)

Alternative Method #1: Salmon River East Fork



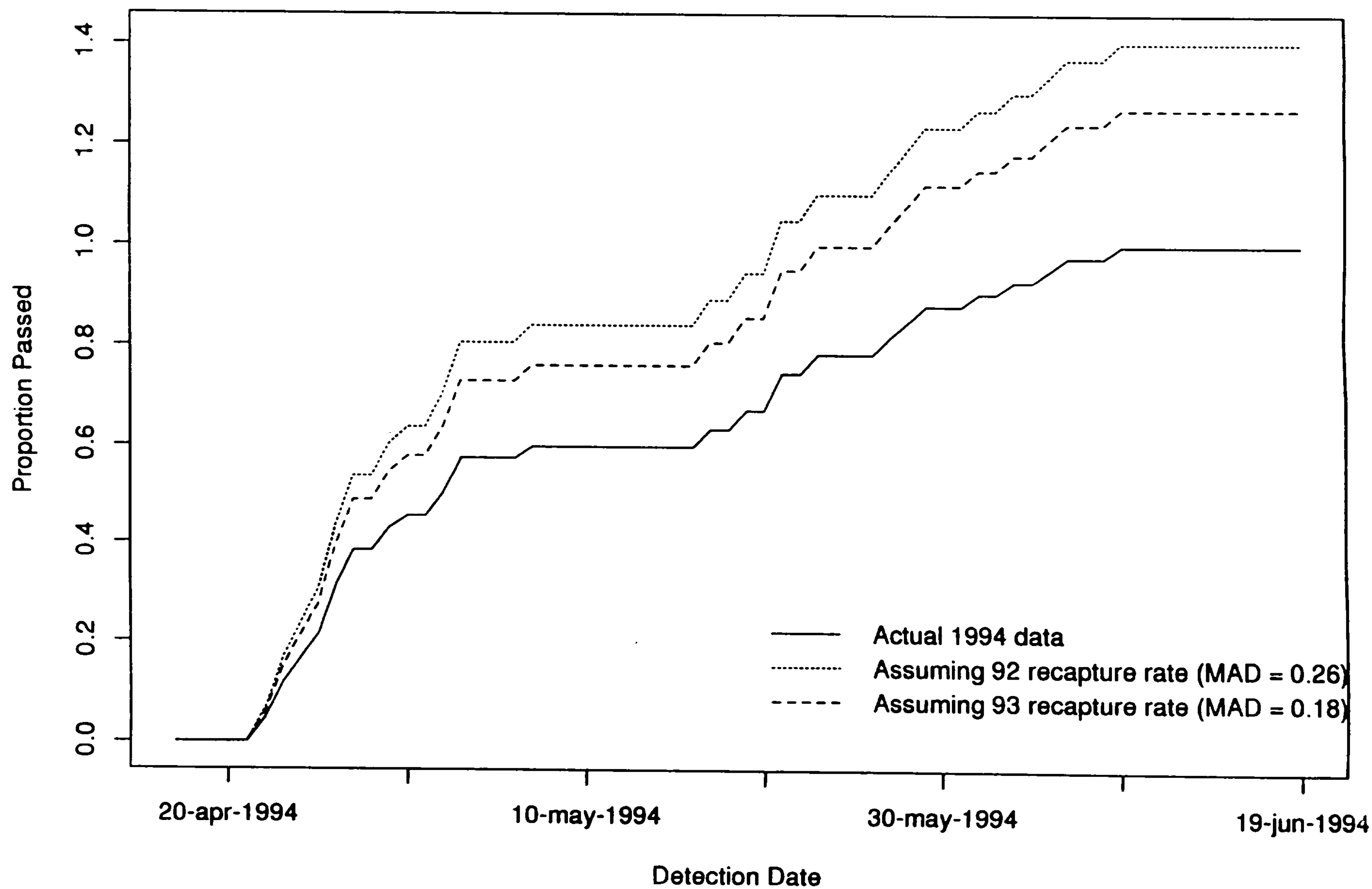
Alternative Method #1 : Salmon River South Fork





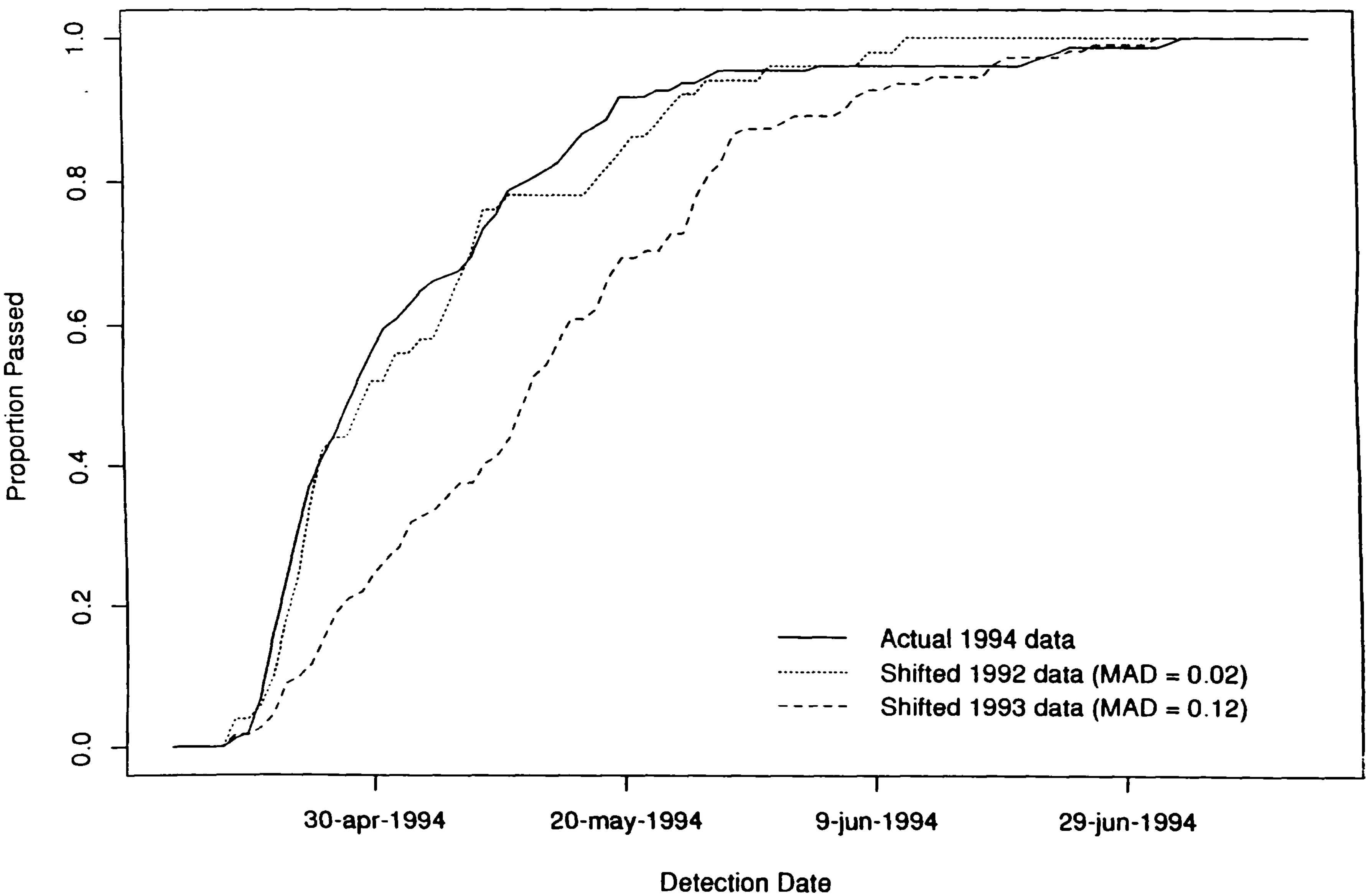
Alternative Method #1: Secesh River

Alternative Method #1: Valley Creek

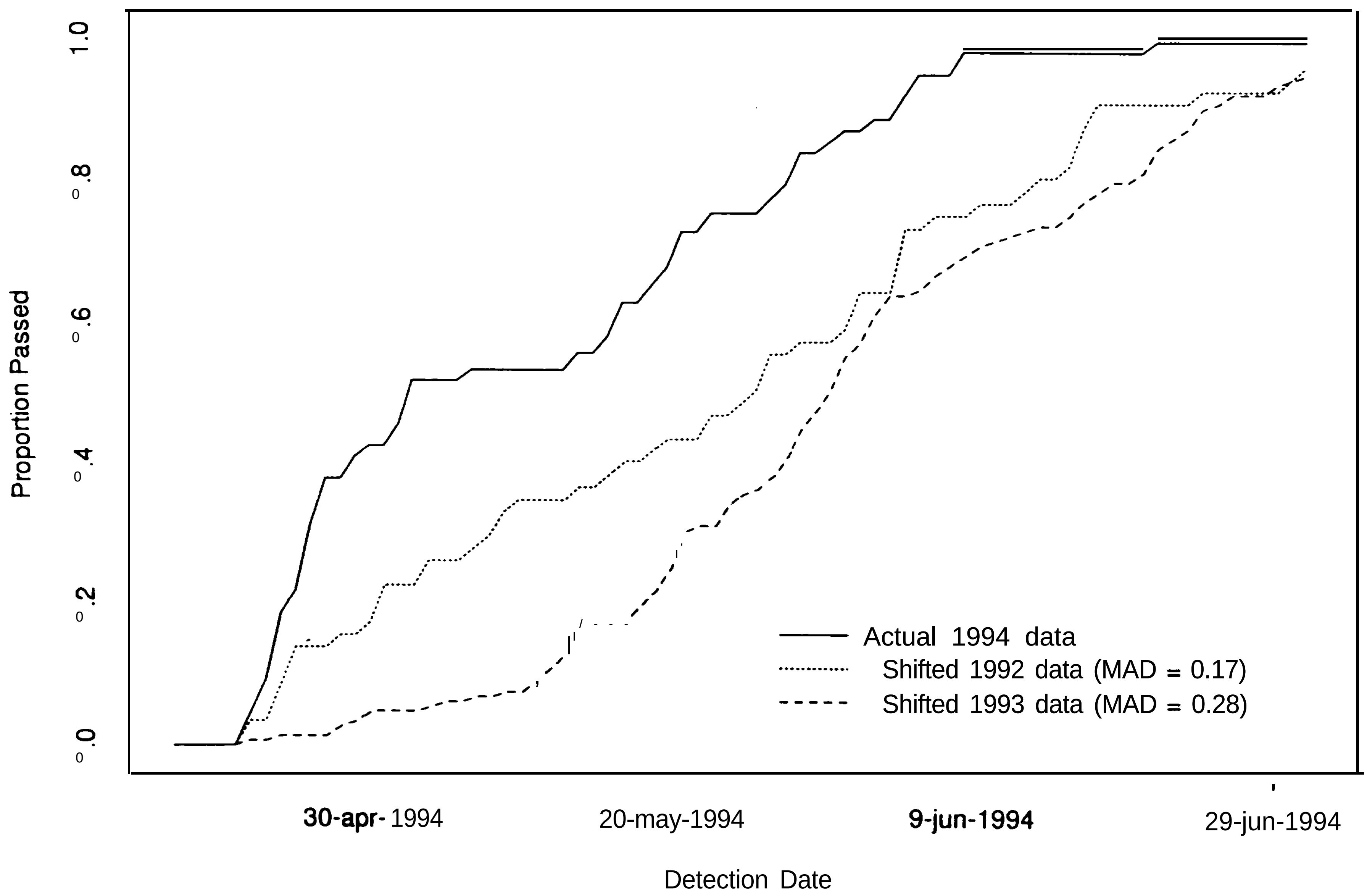


Alternative Method #2 Plots

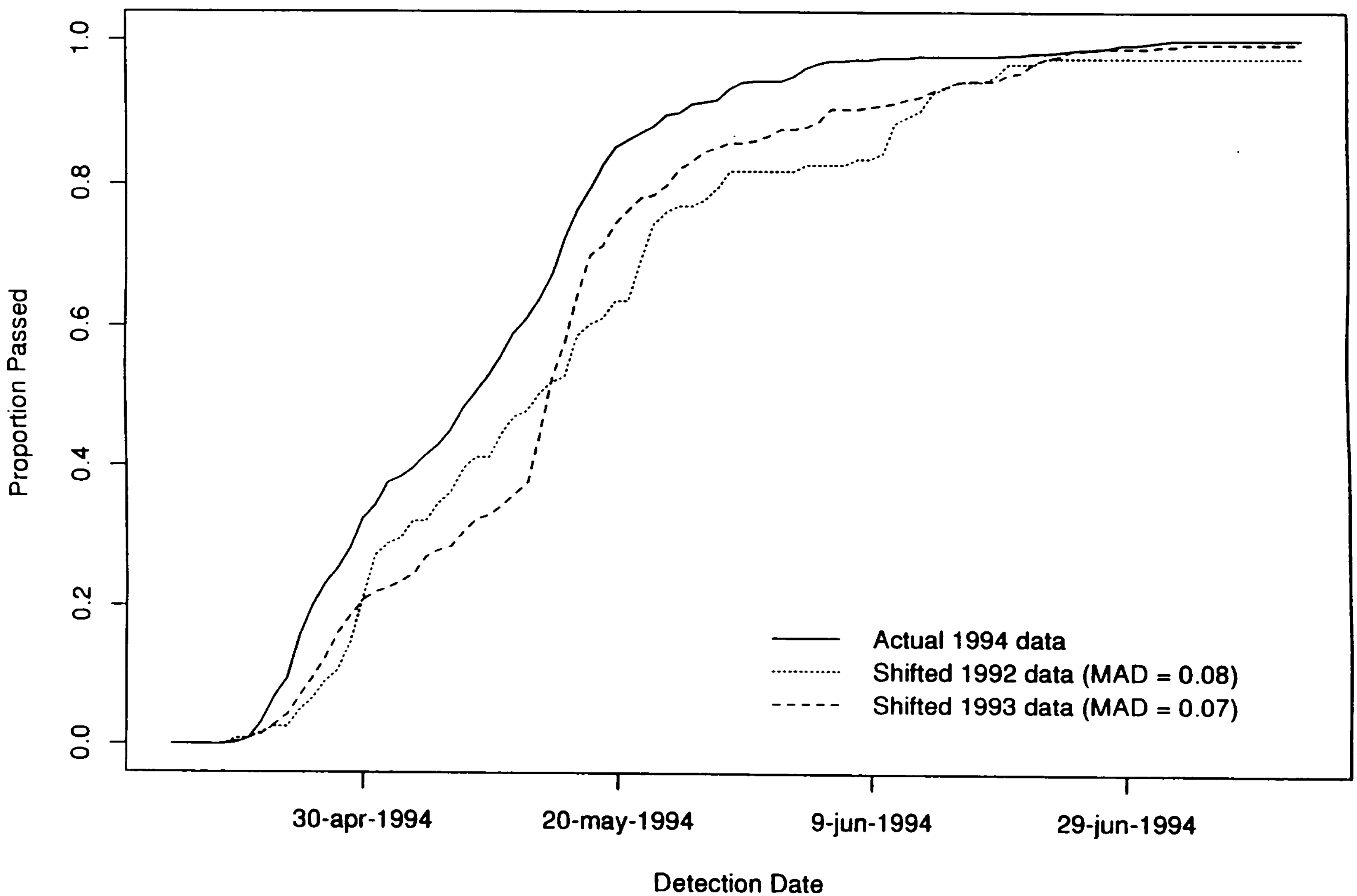
Alternative Method #2: Pahsimeroi



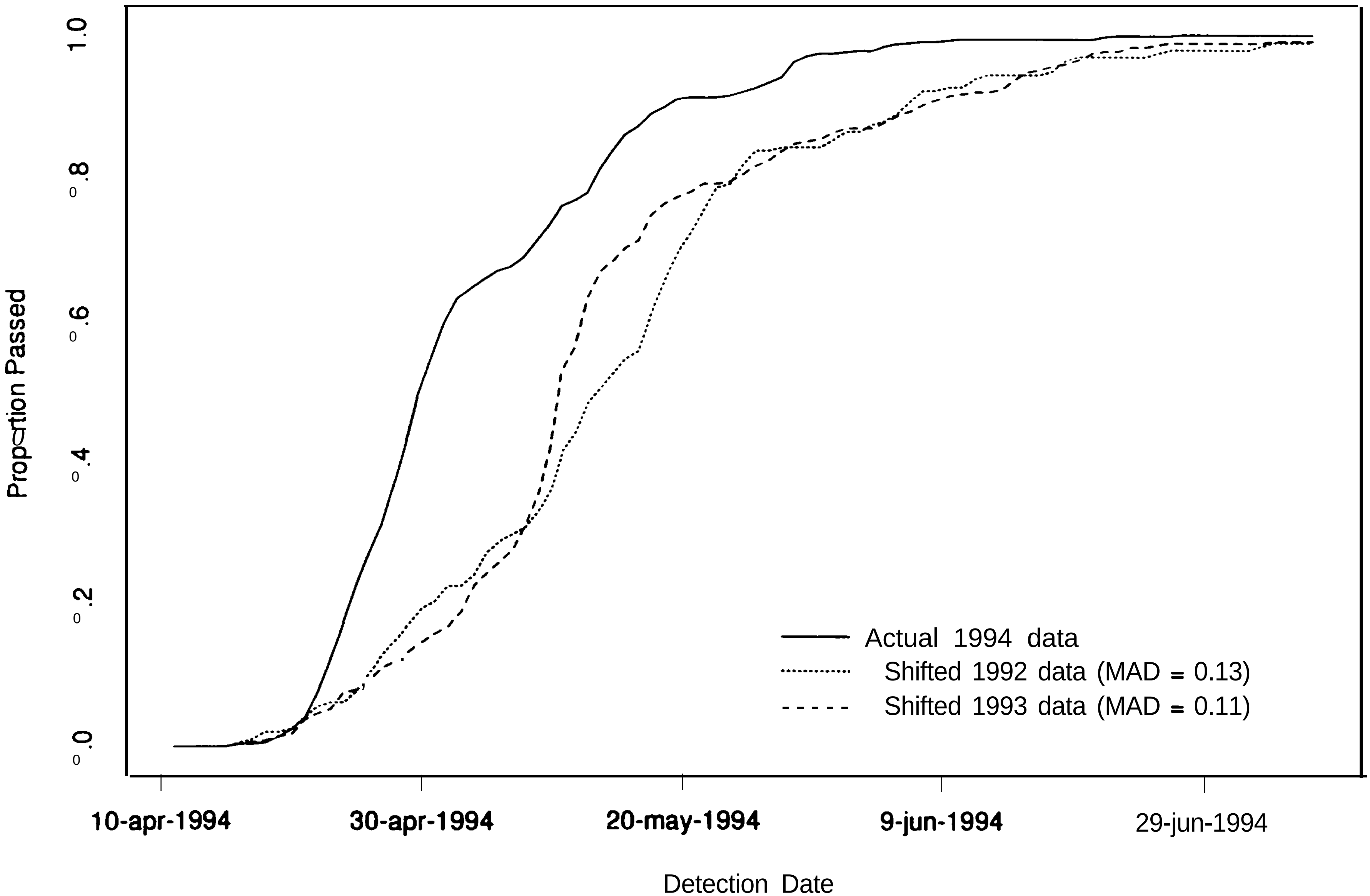
Alternative Method #2: Upper Salmon



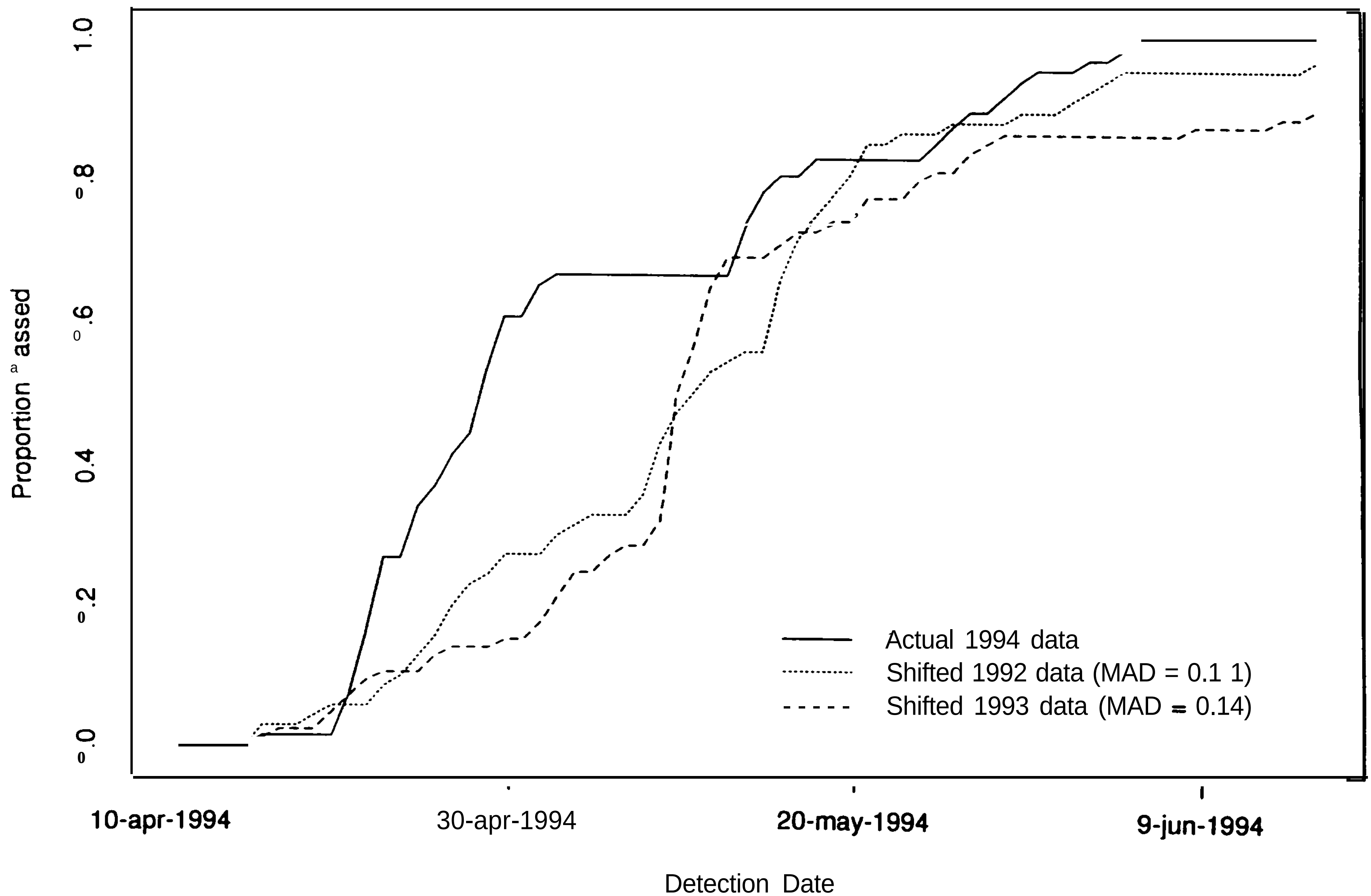
Alternative Method #2: South Fork Salmon



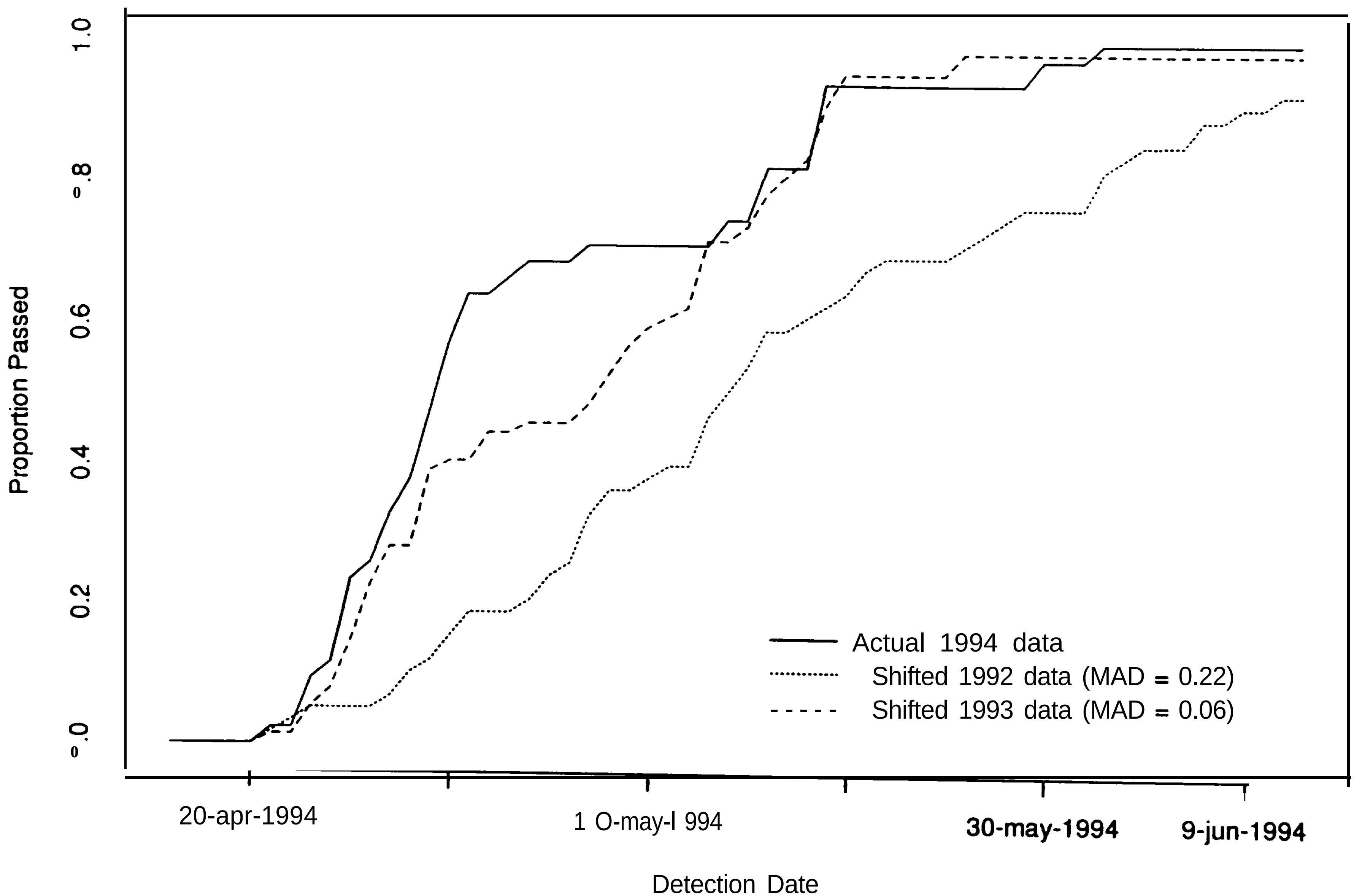
Alternative Method #2: Upper Middle Fork Salmon



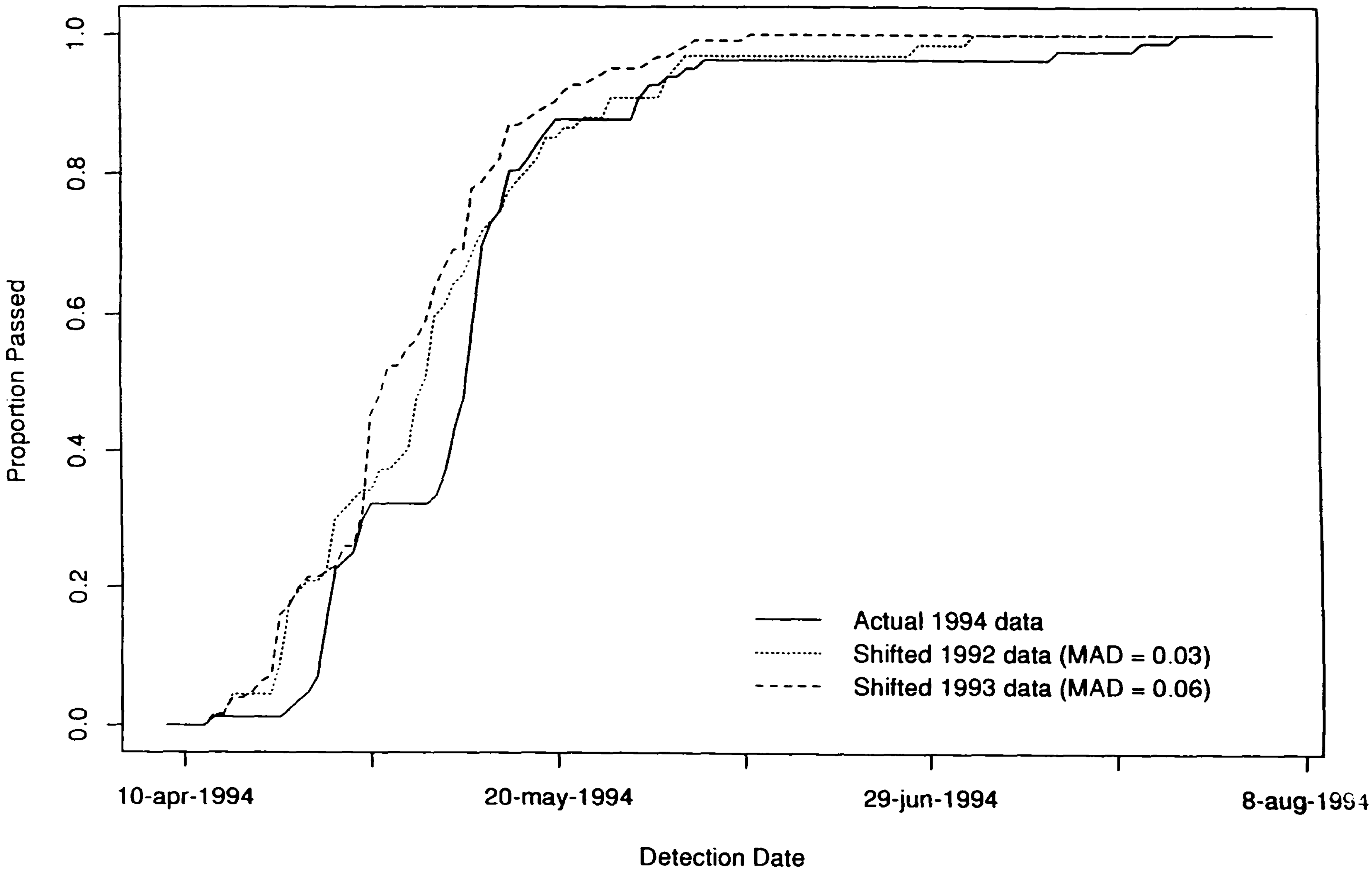
Alternative Method #2: Bear Valley Creek



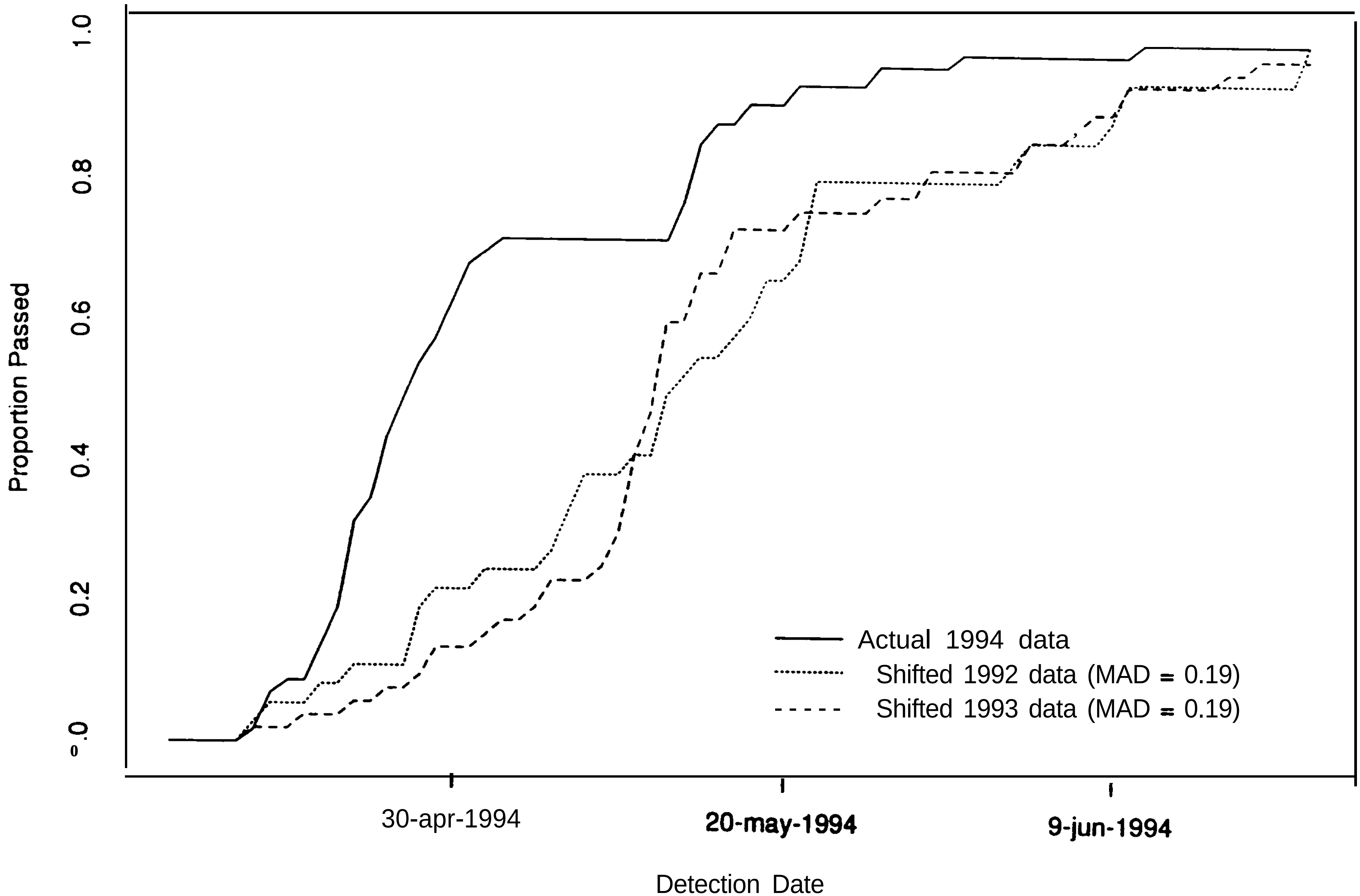
Alternative. Method #2: Big Creek



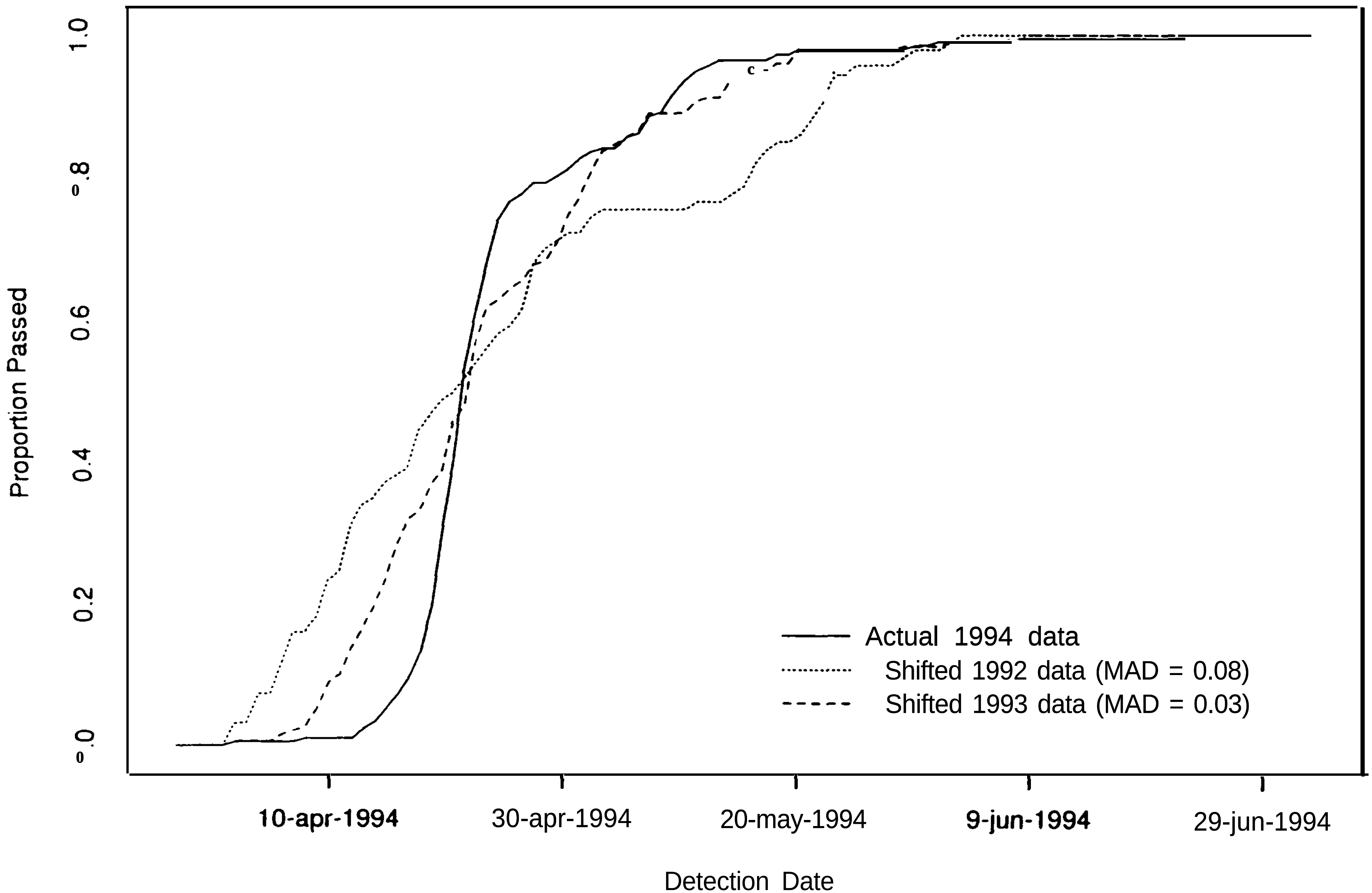
Alternative Method #2: Catherine Creek



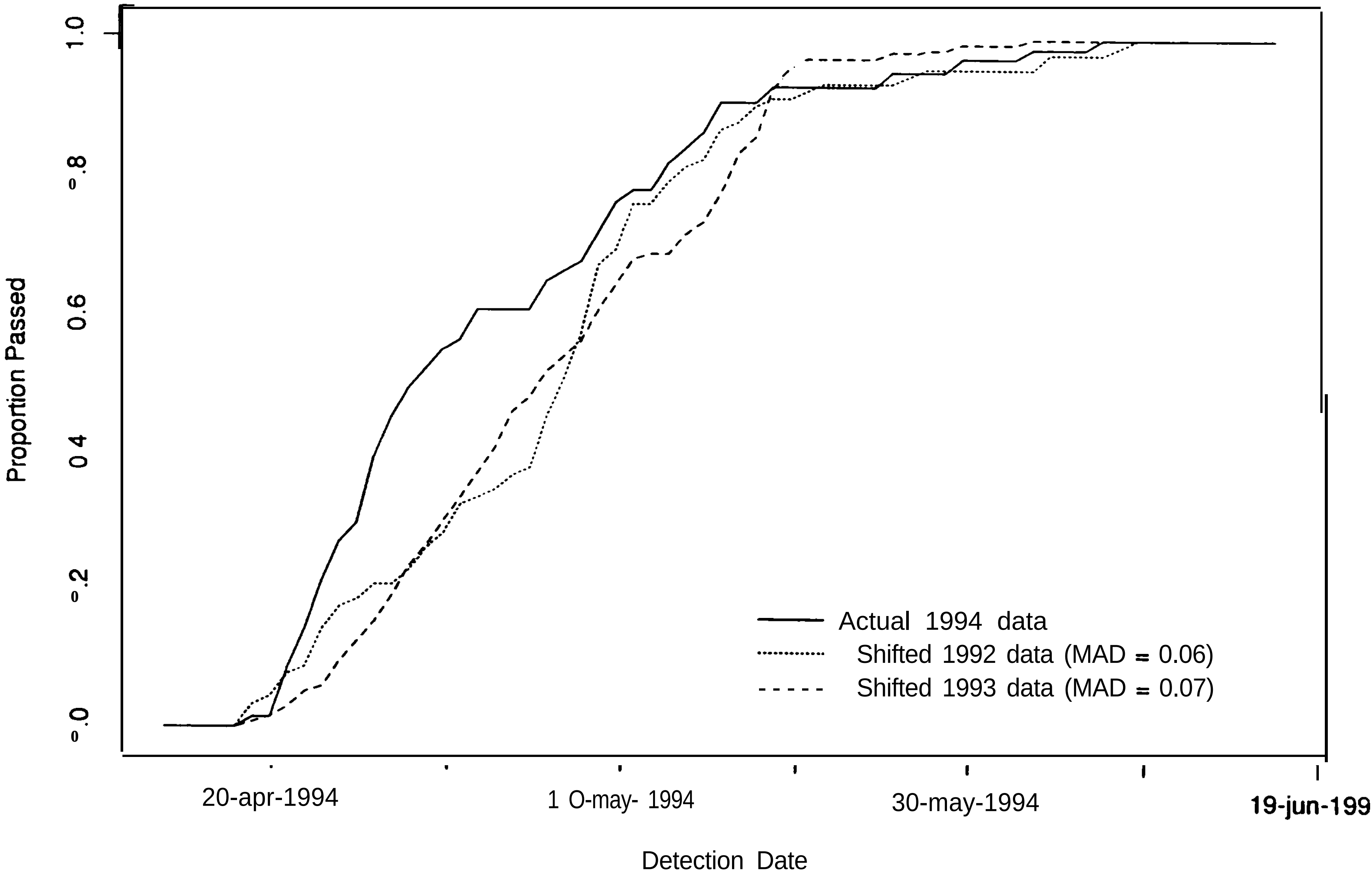
Alternative Method #2: Elk Creek

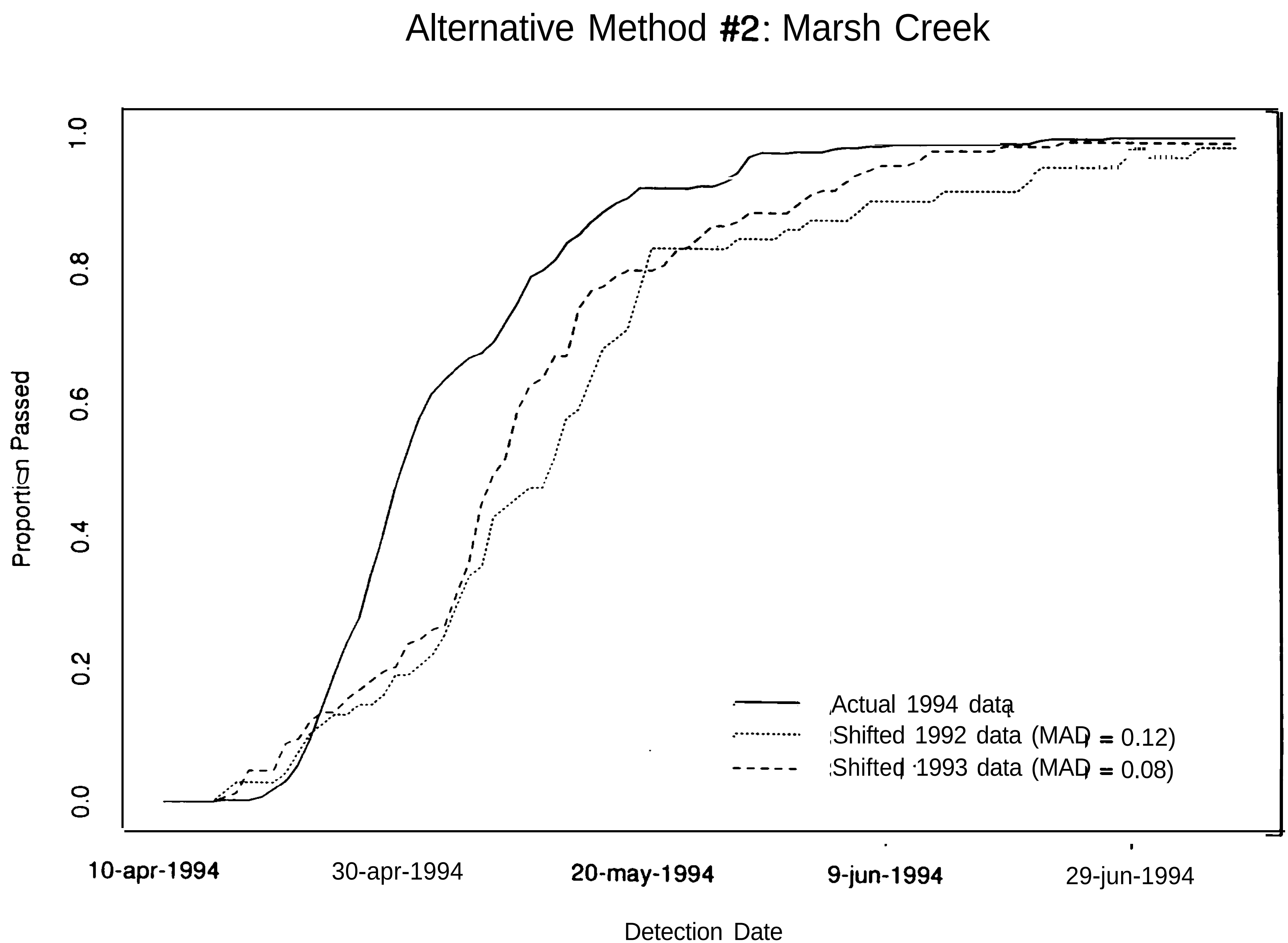


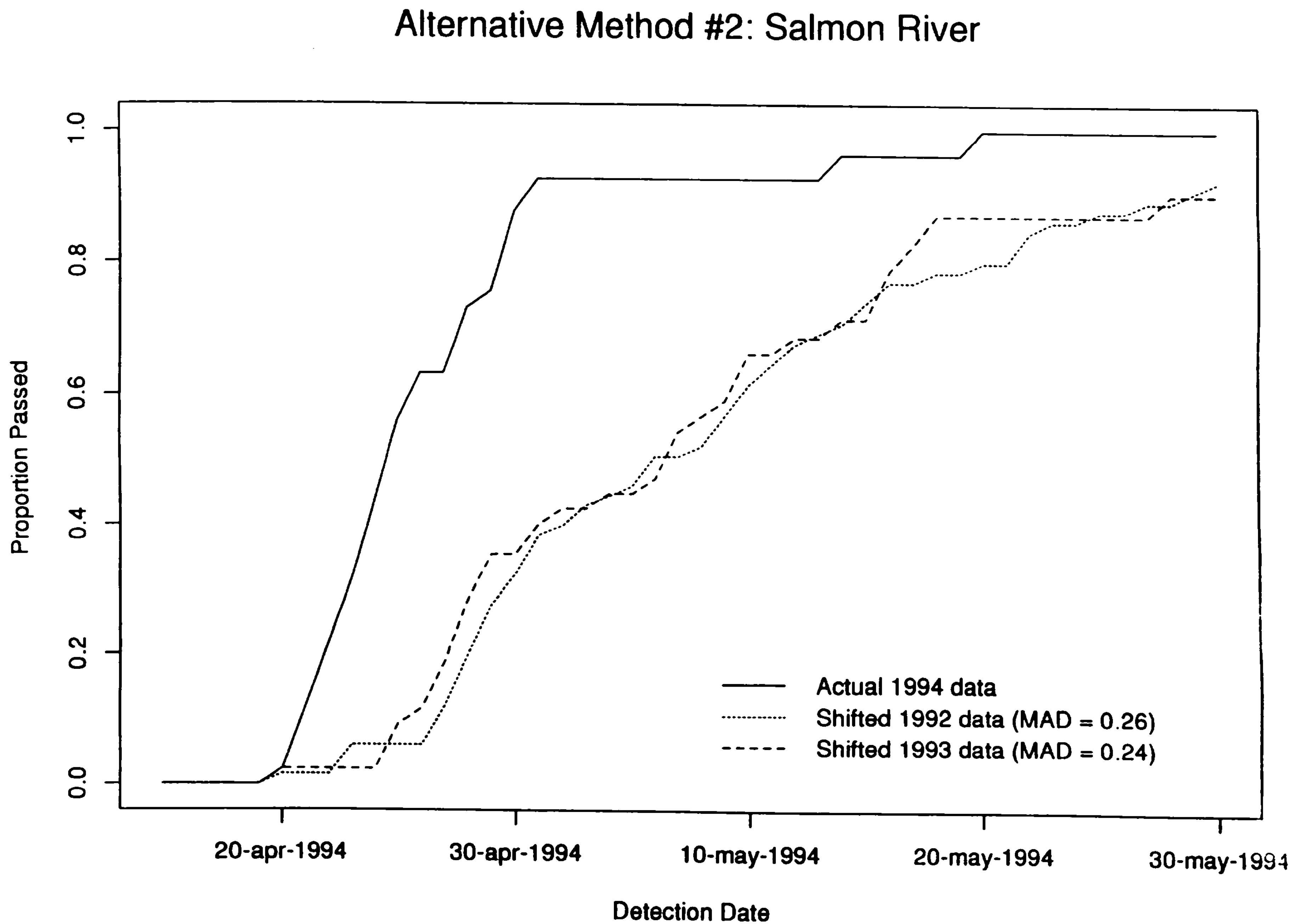
Alternative Method #2: Imnaha River



Alternative Method #2: Lostine River

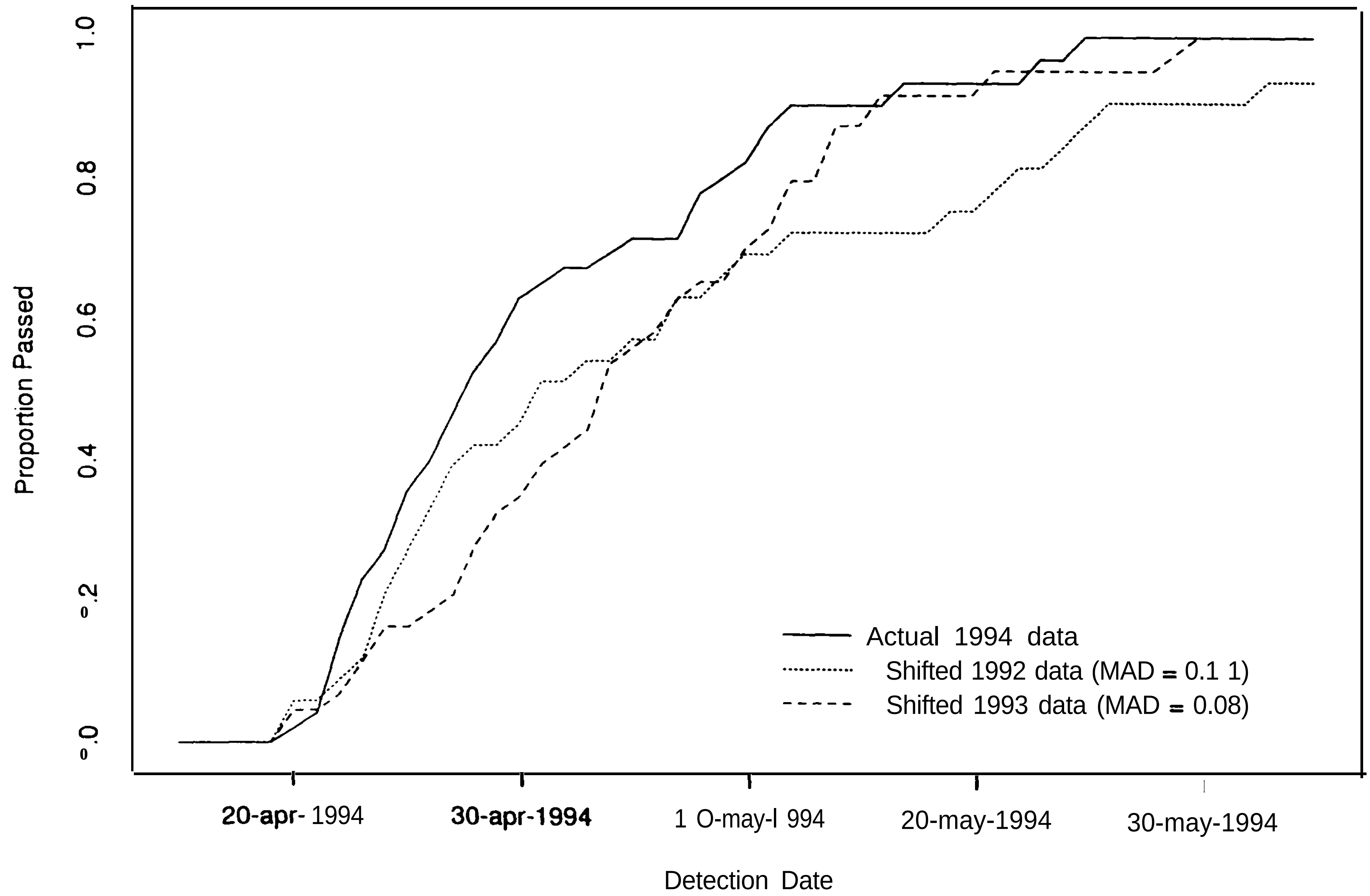



Alternative Method #2: Marsh Creek
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
Actual 1994 data
Shifted 1992 data (MAD = 0.12)
Shifted 1993 data (MAD = 0.08)
10-apr-1994
30-apr-1994
20-may-1994
9-jun-1994
29-jun-1994
Detection Date

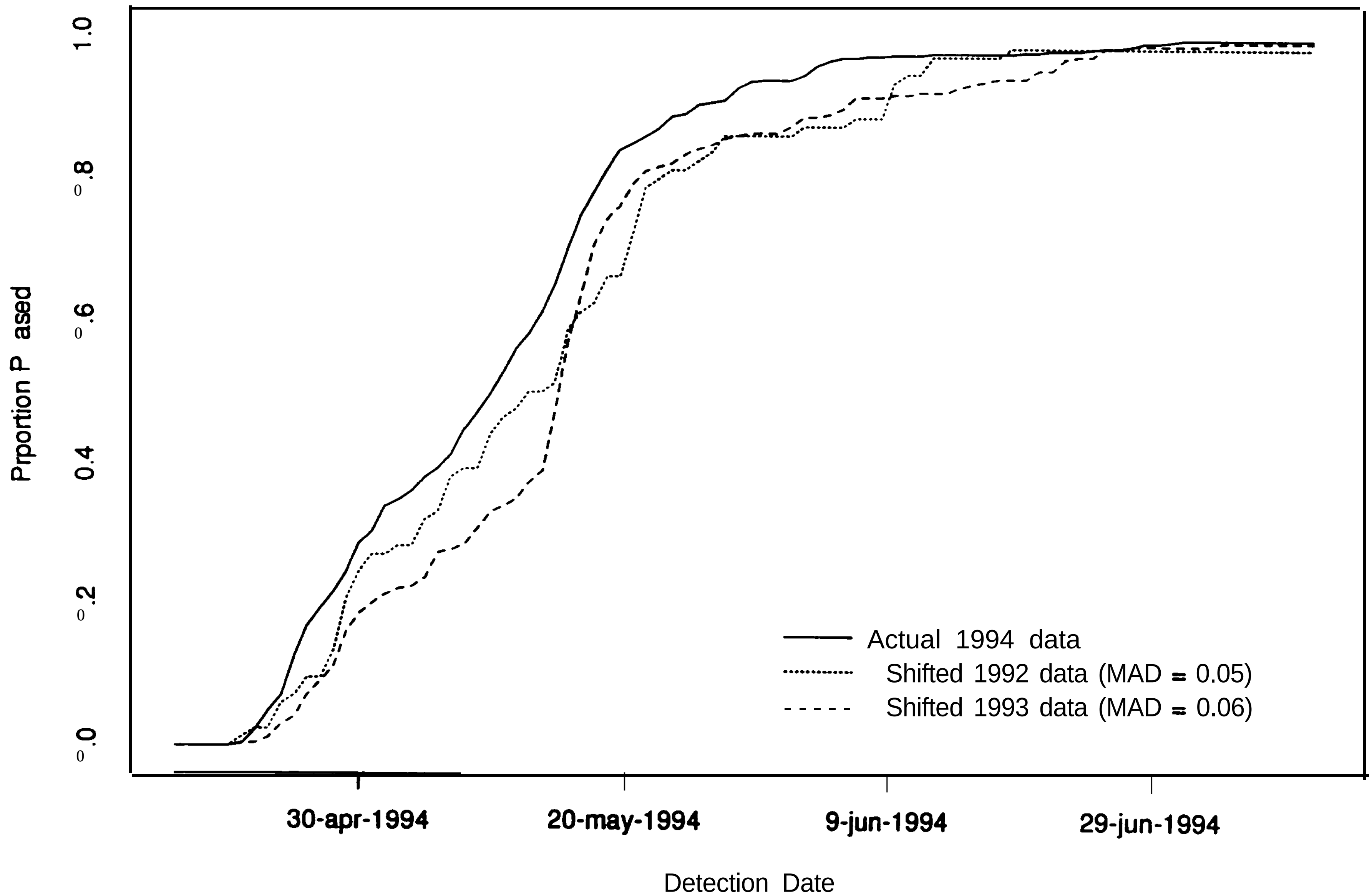


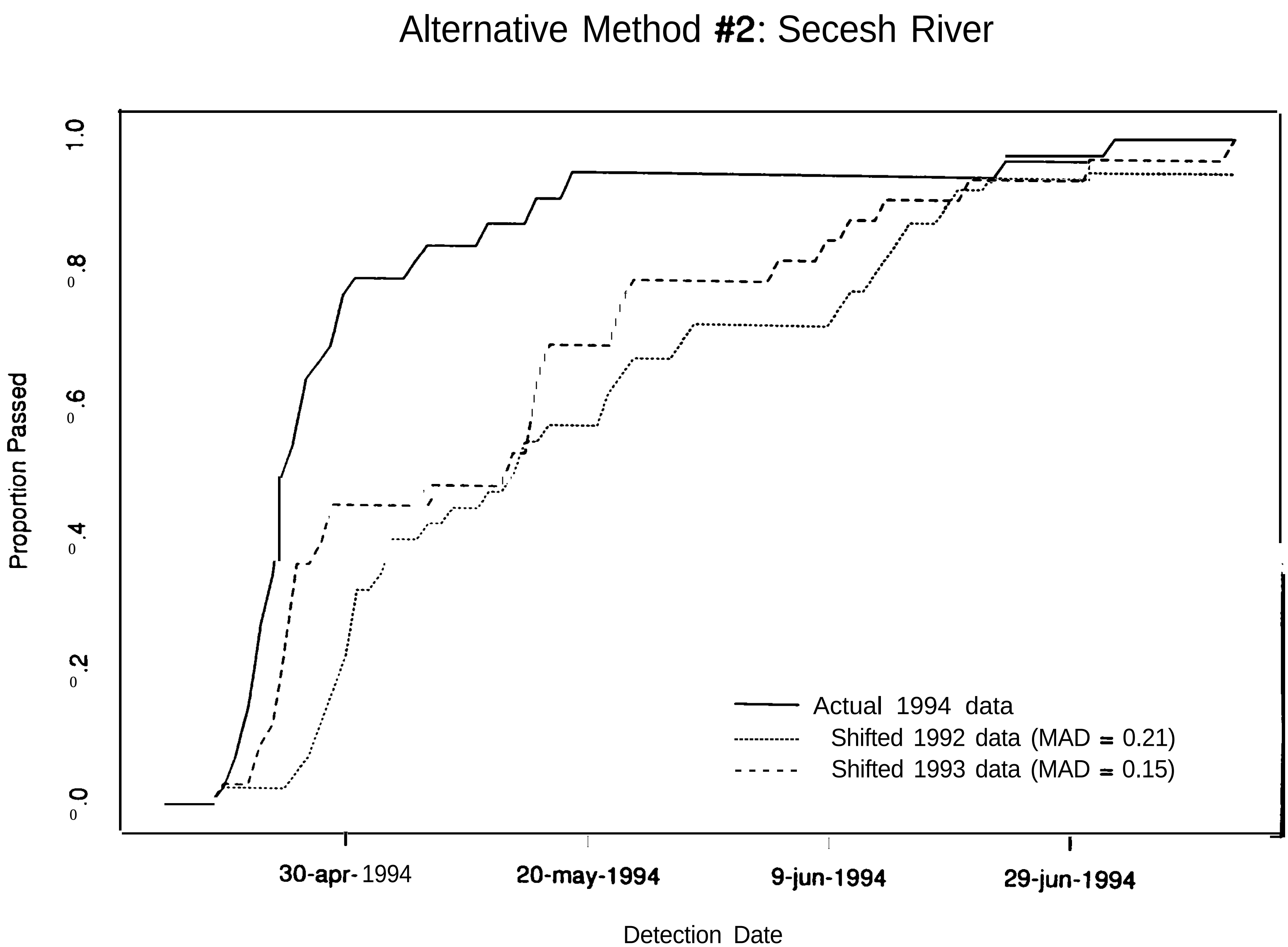
Alternative Method #2: Salmon River
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
20-apr-1994
30-apr-1994
10-may-1994
20-may-1994
30-may-1994
Detection Date
Actual 1994 data
Shifted 1992 data (MAD = 0.26)
Shifted 1993 data (MAD = 0.24)

Alternative Method #2: Salmon River East Fork

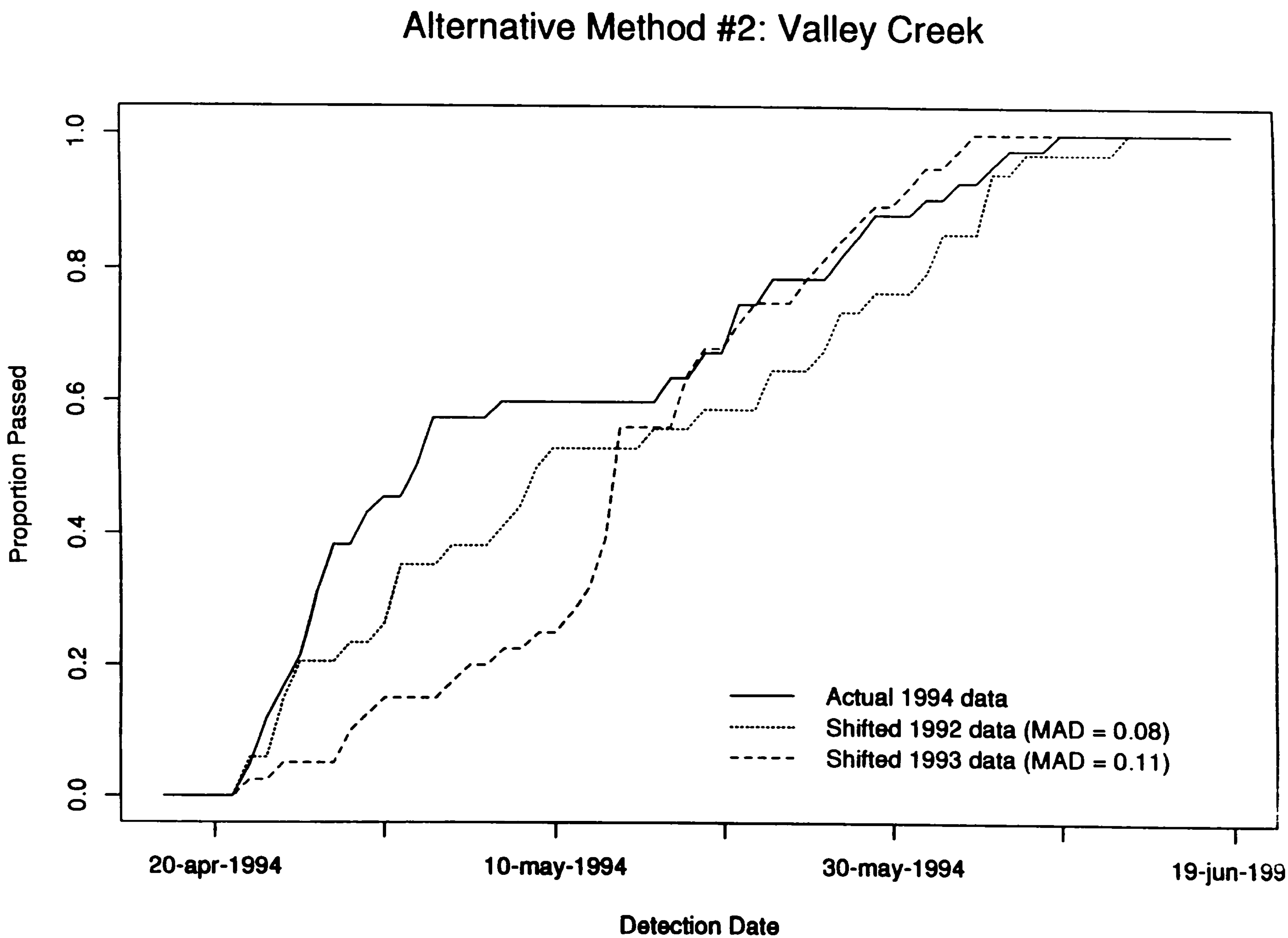


Alternative Method #2: Salmon River South Fork





Alternative Method #2: Secesh River
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
30-apr-1994
20-may-1994
9-jun-1994
29-jun-1994
Detection Date
Actual 1994 data
Shifted 1992 data (MAD = 0.21)
Shifted 1993 data (MAD = 0.15)


Alternative Method #2: Valley Creek
Proportion Passed
0.0
0.2
0.4
0.6
0.8
1.0
20-apr-1994
10-may-1994
30-may-1994
19-jun-1994
Detection Date
Actual 1994 data
Shifted 1992 data (MAD = 0.08)
Shifted 1993 data (MAD = 0.11)

Appendix B

Instructions on the Use of the Forecasting Software

Realtime

RUNS:

The Realtime PIT Forecaster has two operating modes: windows or batch. To run under windows, type "rt" at system prompt:

% rt

The realtime base frame will come up. See following sections for a full description of the various frames.

To run in batch mode, type “t-t -batch -option’*, where option is either "LS" or “Sync”:

% rt -batch -LS or
% rt -batch -Sync

This will results in predictions for all tagsites and tagsite aggregates currently included in the realtime database. The results will be written to a file. either rt.ls.out or rt.sync.out. See section “Batch Output’*.

DATA:

Realtime uses two types of datafiles, flow data files and files with detections of pit tagged fish. We get the detection files from PITAGIS, and the flow files from the Army Corps of Engineers.

The PITAGIS files will have a .csv on the end (comma separated variables) and the ACOE flow files will be of the form rp92xxxx.txt where xxxx is the month and day (i.e. 0424).

These files are parsed and sorted to become the Realtime data files (.flow and .obs) which are kept in the “data” directory.

New data files can be placed in the “newdata” directory. Realtime will automatically read any .txt of csv file from the “newdata” directory, parse the files and then append them to the existing data files, if the files are of the proper format. Once a new data file has been read. a “done” is added to the end of it to let Realtime now the file has already been read.

Batch Output

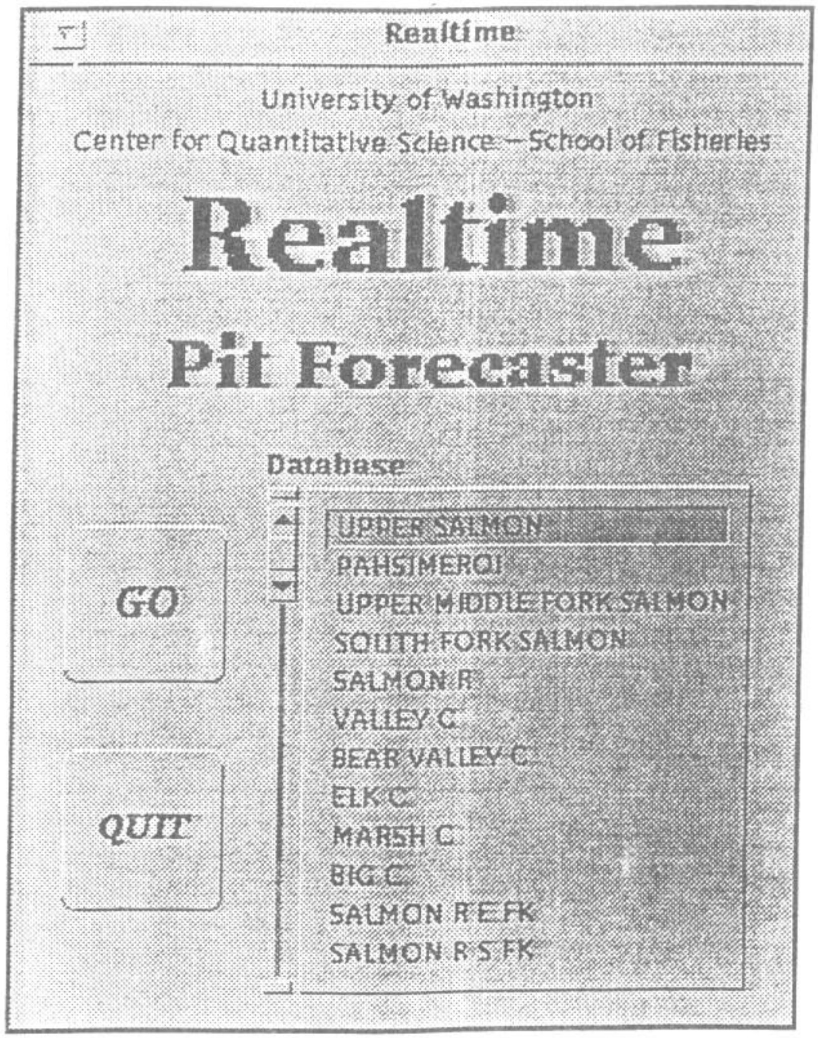
PIT FORECASTER: forecast summary

Wed Apr 27 14:35:01 1994

Tag Site	Days into Run	Fish so Far	Ave Hist Perc	Pred Perc	MAD	Days till 75%	Days till 85%	Days till 95%
UPPER SALMON	5	12	15	12	6.00	37	43	54
PAHS IMEROI	8	45	35	31	13.67	17	24	39
UPPER MIDDLE FORK SALMON	11	96	21	16	11.00	22	32	45
SOUTH FORK SALMON	7	47	37	24	2.00	23	41	56
SALMON R	7	18	25	23	15.00	18	24	38
VALLEY C	7	20	20	16	8.40	32	37	44
BEAR VALLEY C	11	18	20	2	0.00	32	43	55
ELK C	9	18	17	11	4.67	25	36	49
MARSH C	11	59	20	15	11.50	23	29	45
BIG C	6	5	14	17	14.00	25	32	42
SALMON R E FK	7	13	38	37	5.67	15	20	34
SALMON R S FK	6	3	23	17	7.25	29	38	51
SECESH R	7	14	43	35	5.60	25	40	60
IMNAHA R	25	136	47	43	12.25	10	18	29
LOSTINE R	19	22	22	28	.33 10	10	15	24
CATHERINE C	4	6	14	23	5.33	19	25	37

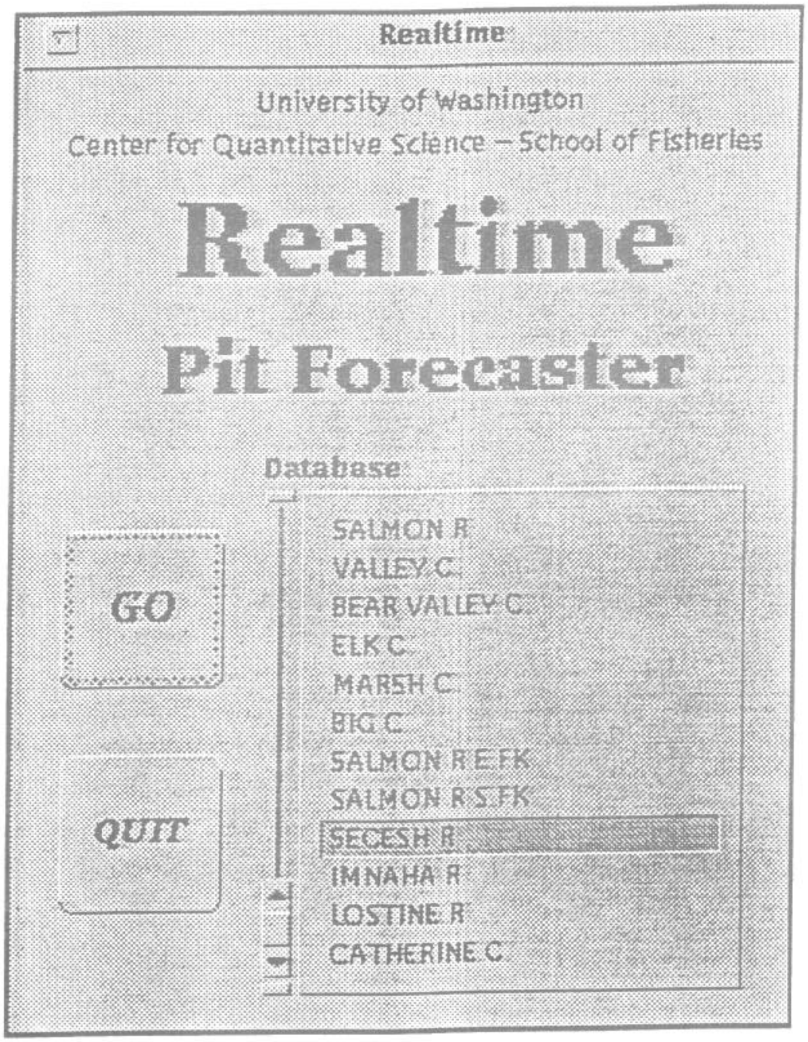
Tag Site:	The first four are aggregates. the rest are individual tag sites.
Days into Run:	The number of days since the first detection.
Fish so Far:	The number of fish detected this year.
Ave Hist Perc:	The average percent passed over historical years on this date.
Pred Perc:	The predicted percentile.
MAD	Mean Absolute Deviation over historical years at this percentile.
Days till 75%	The number of days till we are at the 75th percentile.
Days till 85%	The number of days till we are at the 85th percentile.
Days till 95%	The number of days till we are at the 95th percentile.

Tagsites Frame



Type "rt" at system prompt and the Realtime base frame comes up.

Tagsites Frame

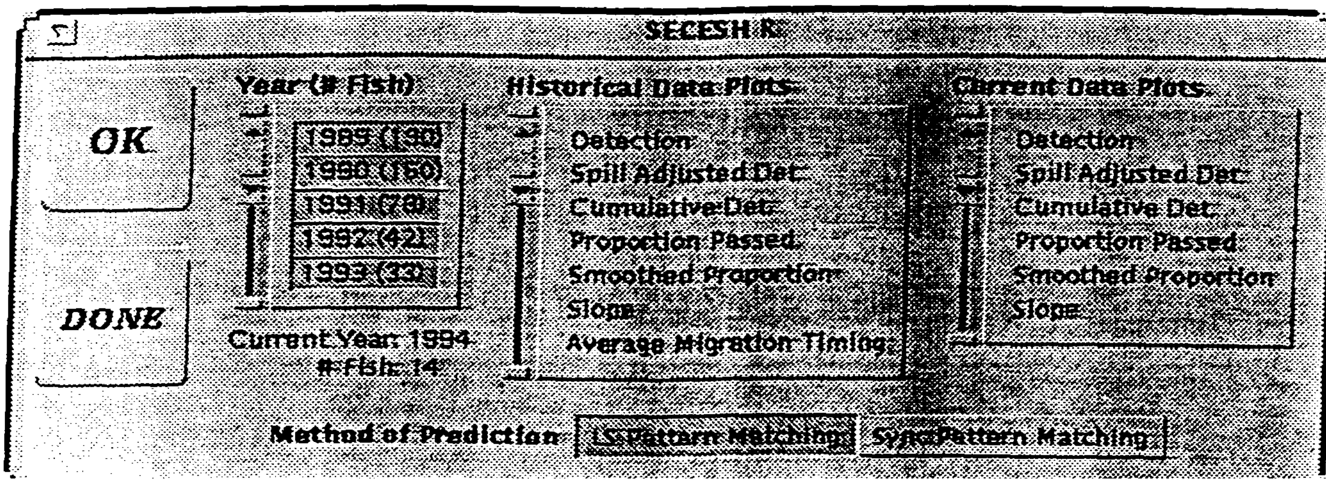


Select database by pointing to named databases with the mouse and clicking the left mouse button. Use scroll bar to see database names that are out of view. To deselect a database, either click again on the name with mouse, or select another database.

These database names are the PIT Tag Information System names for either individual tag sites or in some cases, aggregations of multiple tag sites. The first four names in the menu are aggregates: UPPER SALMON, PAHSIMEROI, UPPER MIDDLE FORK SALMON, and SOUTH FORK SALMON. The rest are individual tag sites.

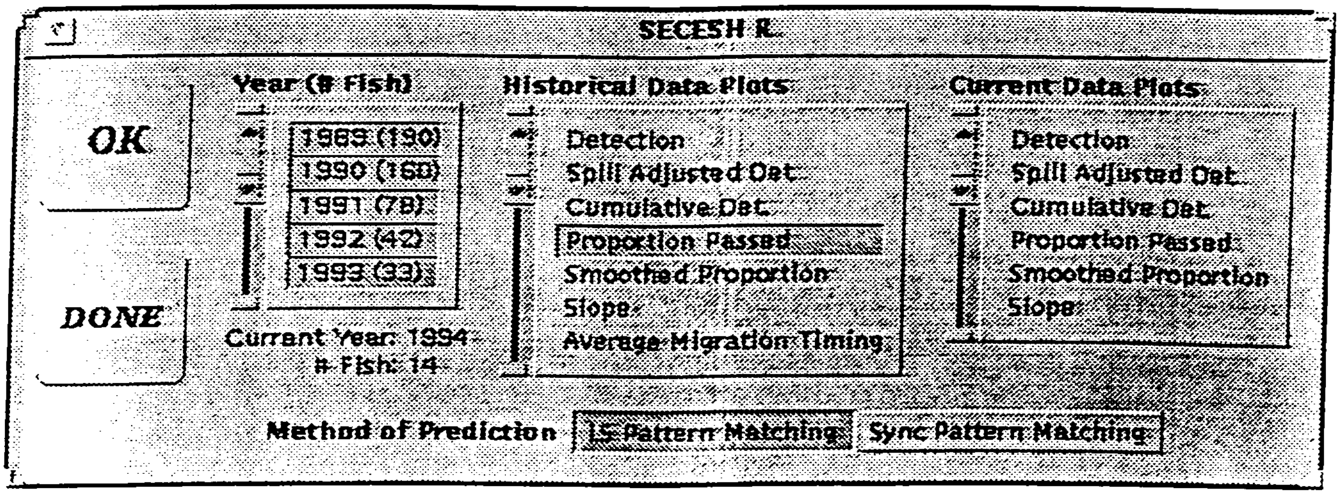
After you have selected your database, click once on "GO" with the mouse button. This brings up the river frame.

River Frame



In the river frame you can look at the data, both “Historical” and “Current”. select which historic years the prediction will be based on, and choose which method of prediction should be used.

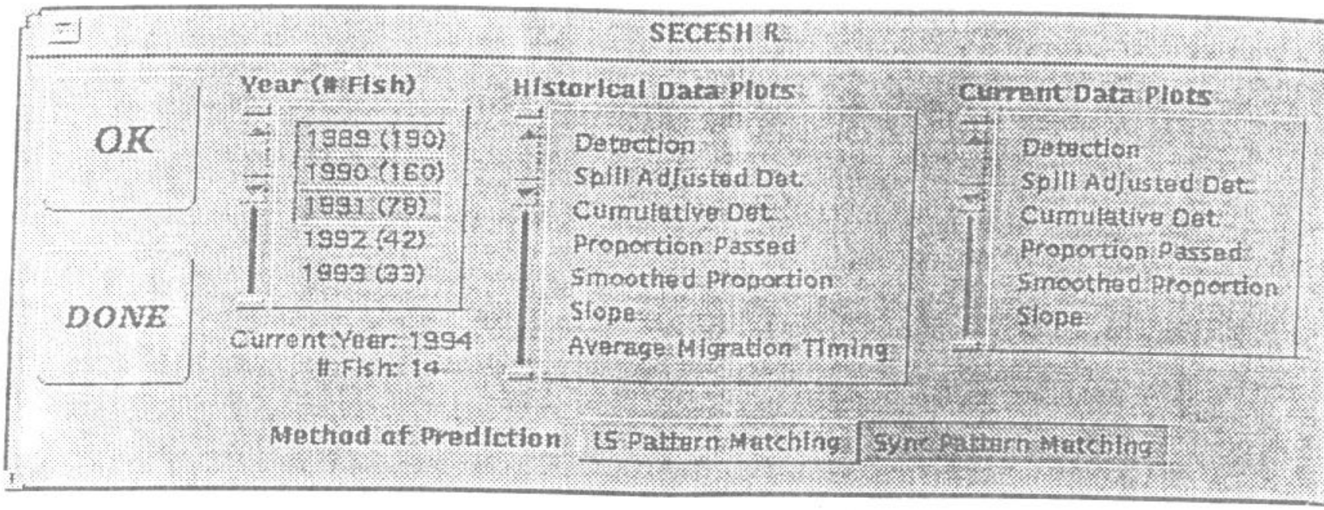
River Frame



Here plots are divided into two sections: "historical'* and "current". For each data set, one year is the current data (generally data from the current year), and the rest is historical data.

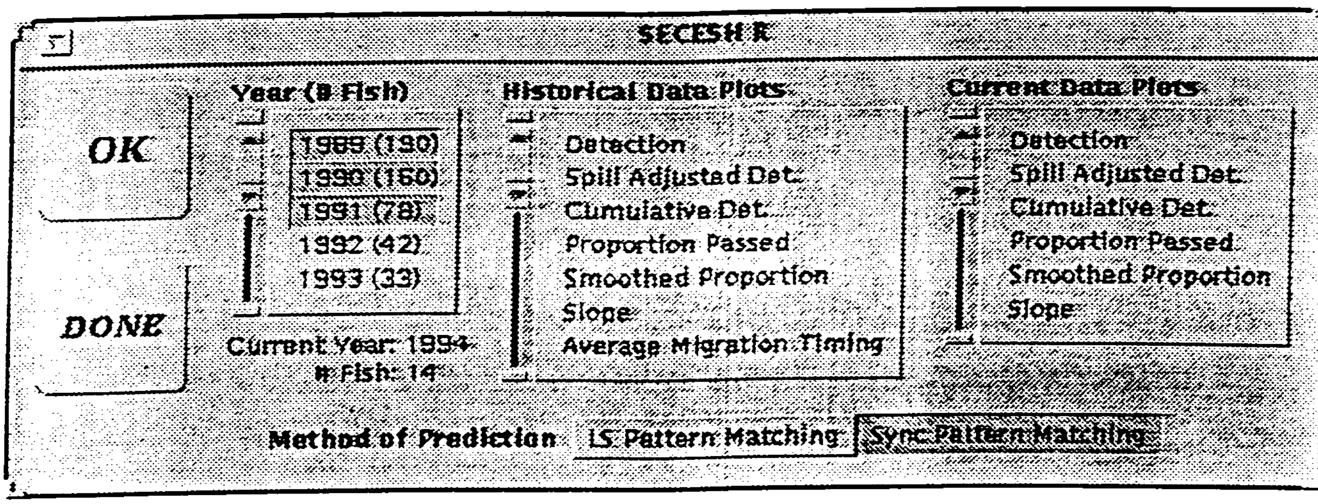
To look at these plots, point and click with mouse button on the plot title in the plot menu. This highlights the plot name and brings up the plot. Clicking on a highlighted plot name destroys that plot. The plots are described in the sections "Historical Data Plots" and "Current Data Plots".

River Frame



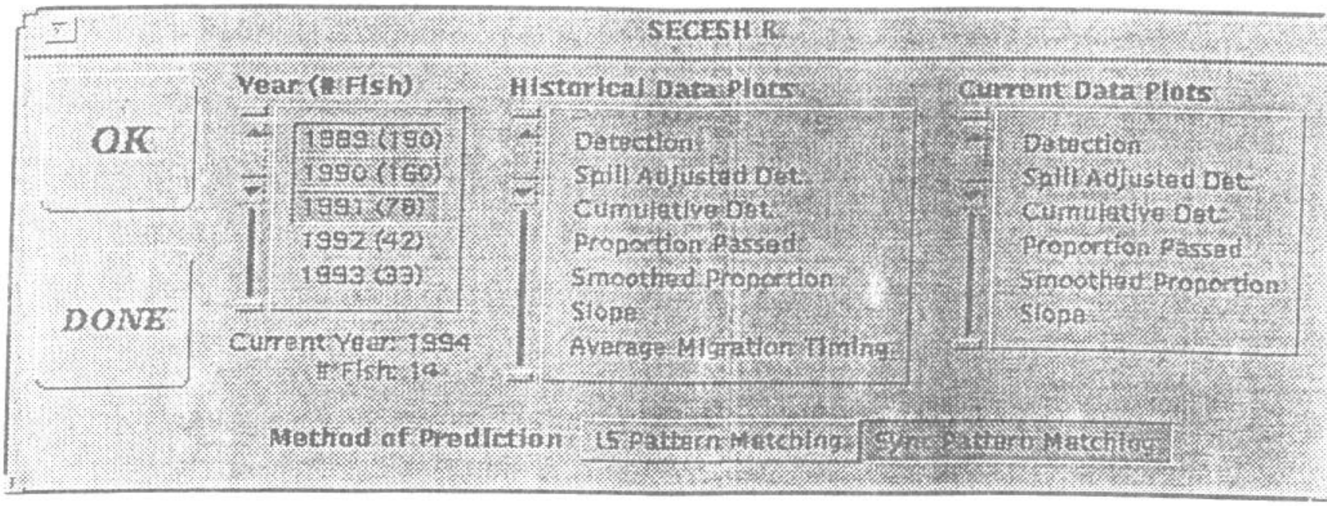
A menu with the title "Year (# Fish)" contains the years and the number of detections for that year. When the river frame comes up, these are all highlighted indicating that all will be used as historical data in the analysis. To remove a year from consideration, click once with the mouse button. Clicking a second time will re-select the year.

River Frame



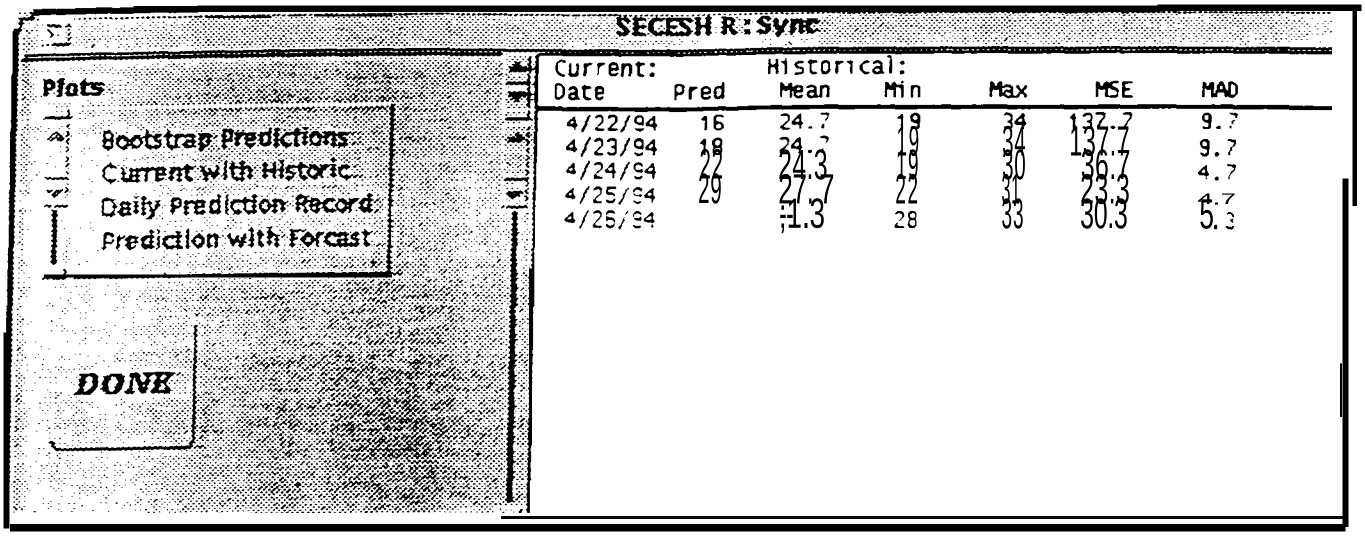
Two methods are implemented currently: "LS Pattern Matching" and "Sync Pattern Matching". Choose the method by clicking once with the mouse button.

River Frame



Click "OK" to begin the prediction. When the results have been computed, the results frame will come up. The "OK" button on the river frame will now be inactive and will remain inactive until you are done with the results frame.

Results Frame



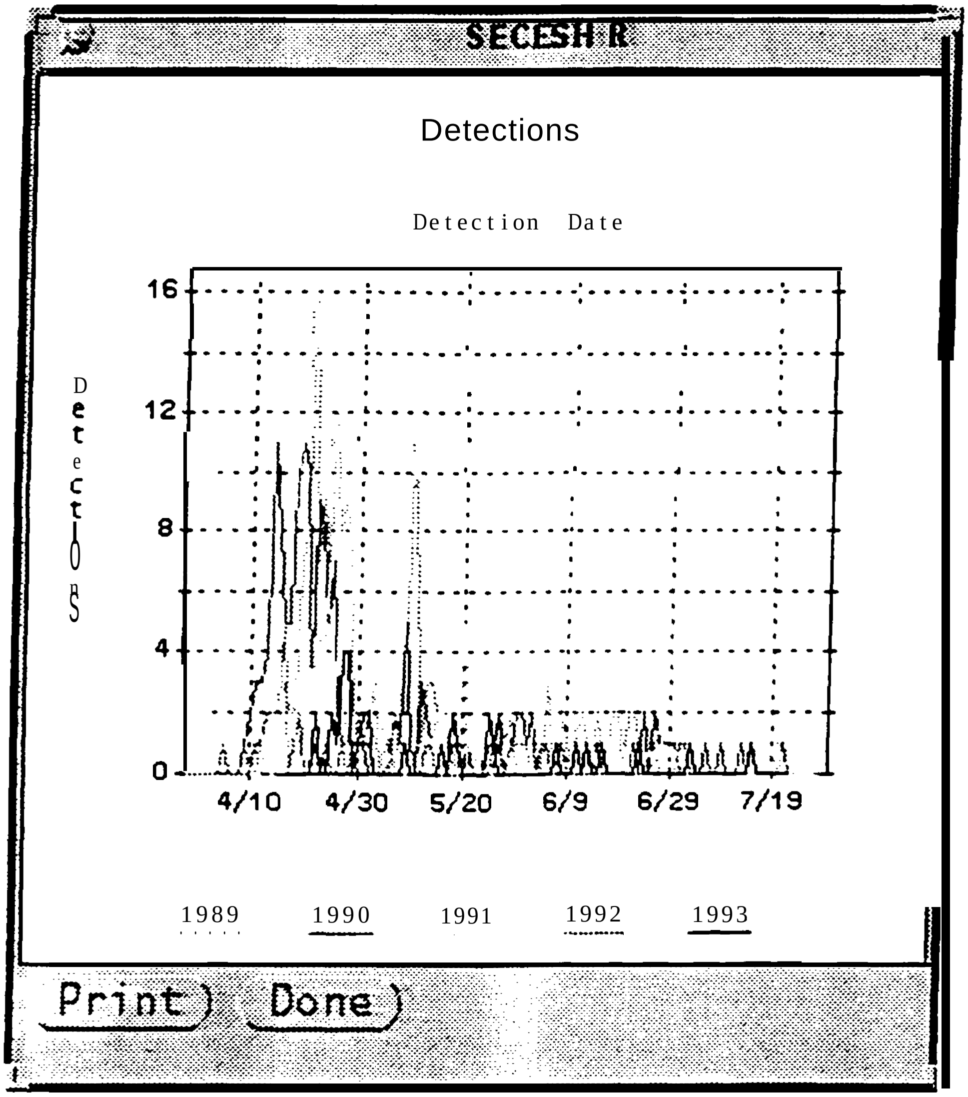
The results frame has two areas: a plot menu and a text window. To look at the various plots, point and click on the plot name. These plots are described in the section “Results Plots”. In the text window is a scrollable table that contains predictions and confidence estimates. Realtime reports the current prediction as well as a record of the predictions made on previous days.

Each row has the date. the prediction and summaries of the predictions performance based on data gained from bootstrap predictions over historic years.

Plots

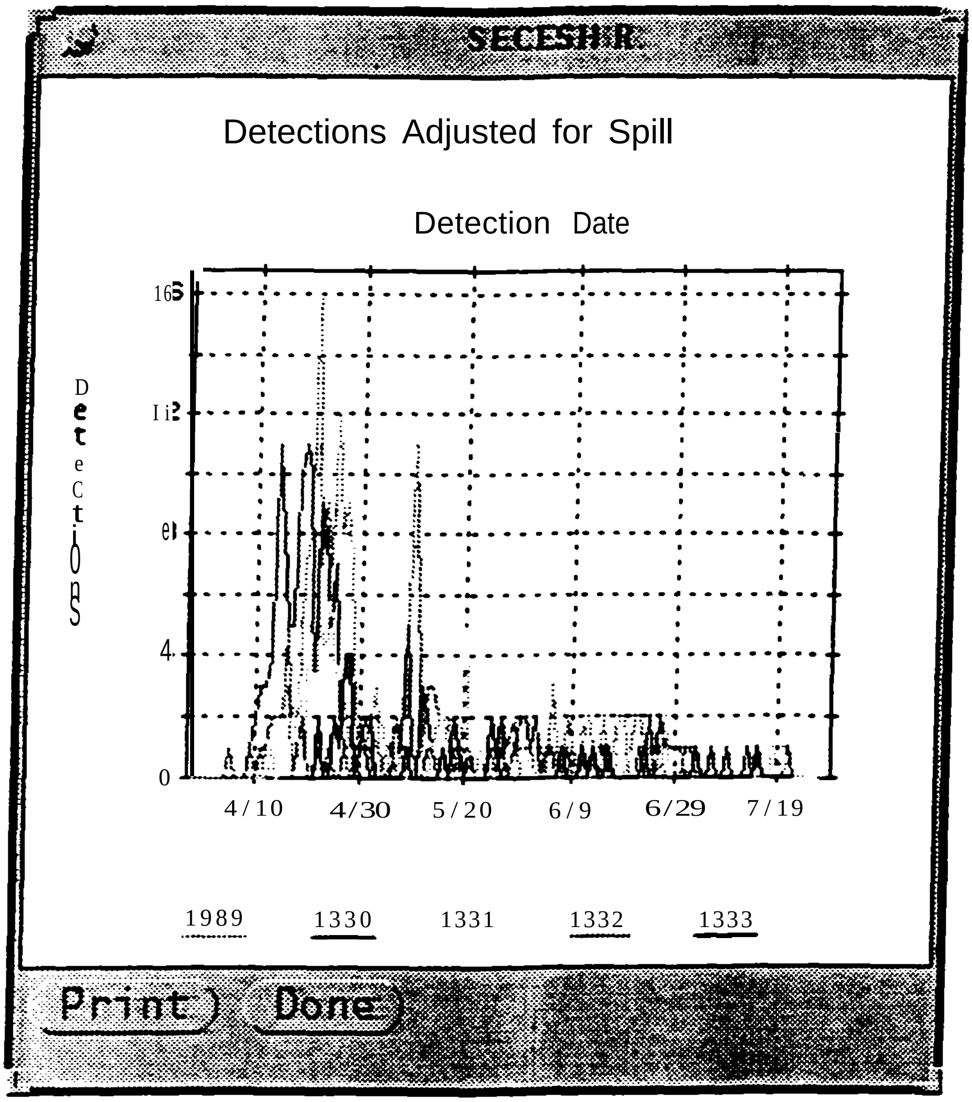
Plots have two buttons to know about: “Print” and “Done”. The print button will send the plot to lpr. The done button will destroy the plot.

Historical Data Plots



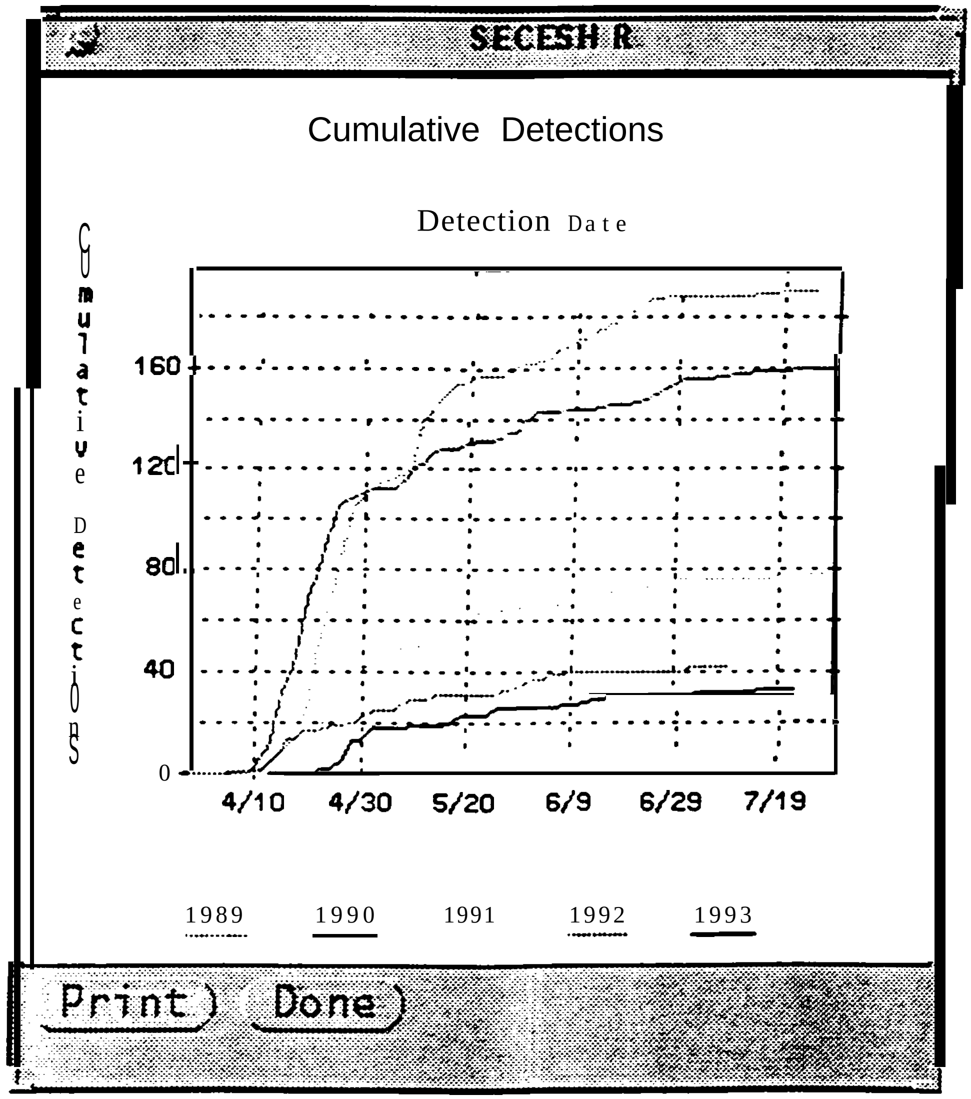
Number of detections vs. detection date

Historical Data Plots



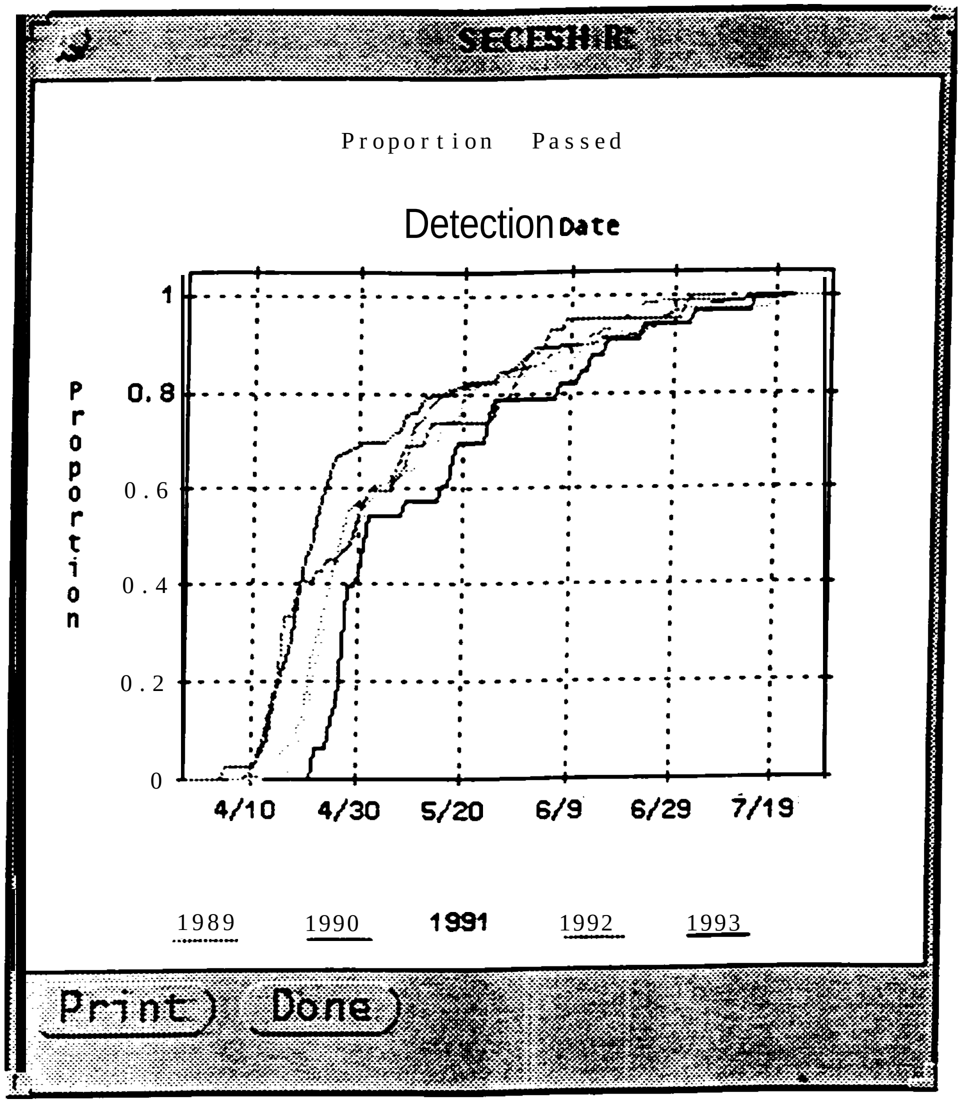
Number of detections adjusted for spill efficiency vs. detection date.

Historical Data Plots



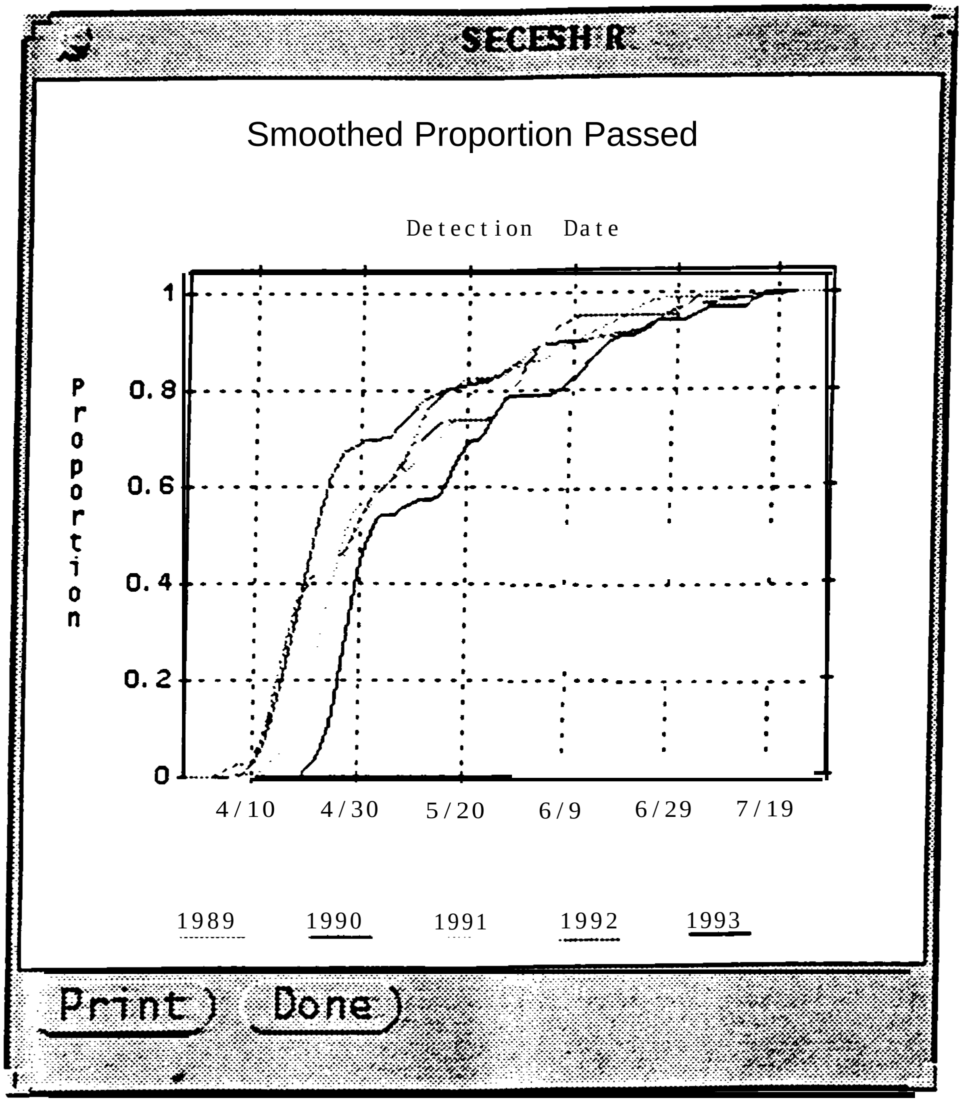
Cumulative detections (adjusted for spill) vs. detection date.

Historical Data Plots



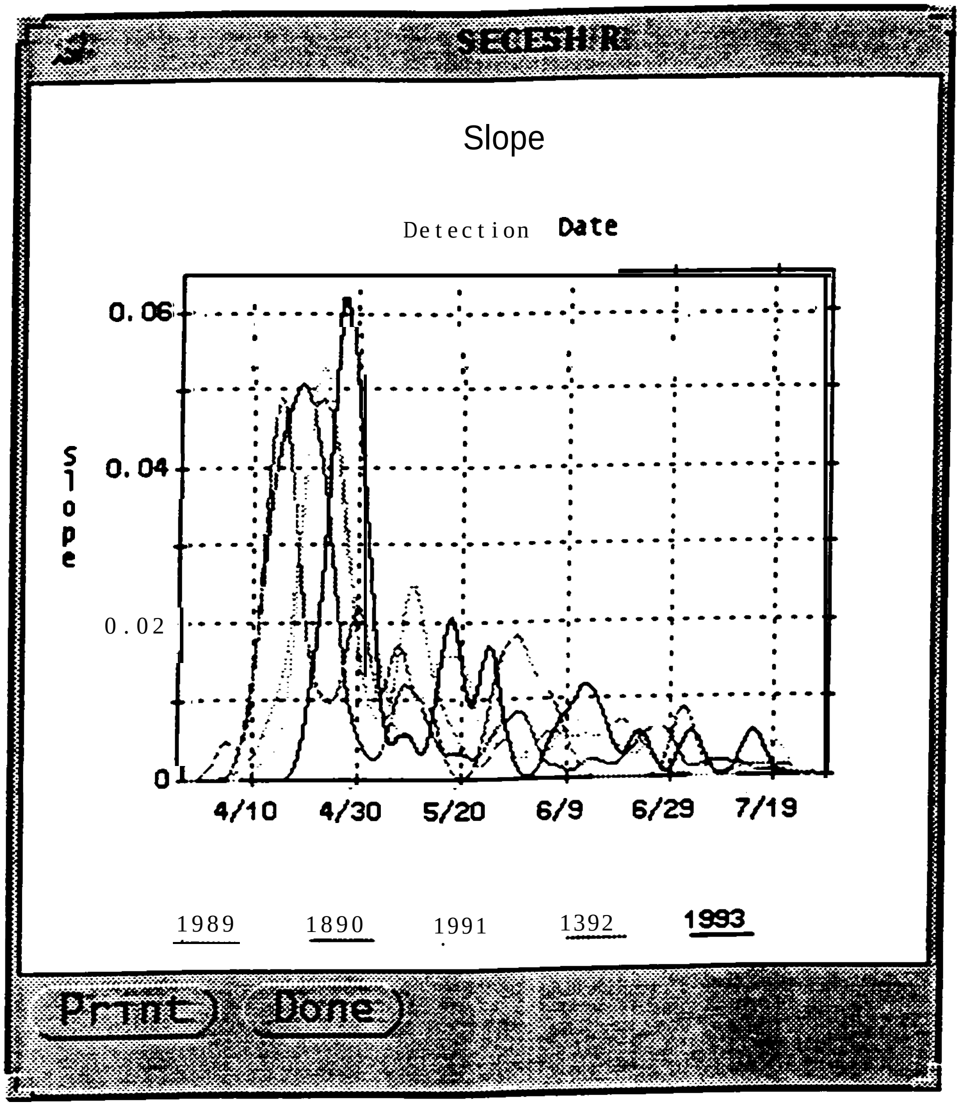
Cumulative detections normalized vs. detection date.

Historical Data Plots



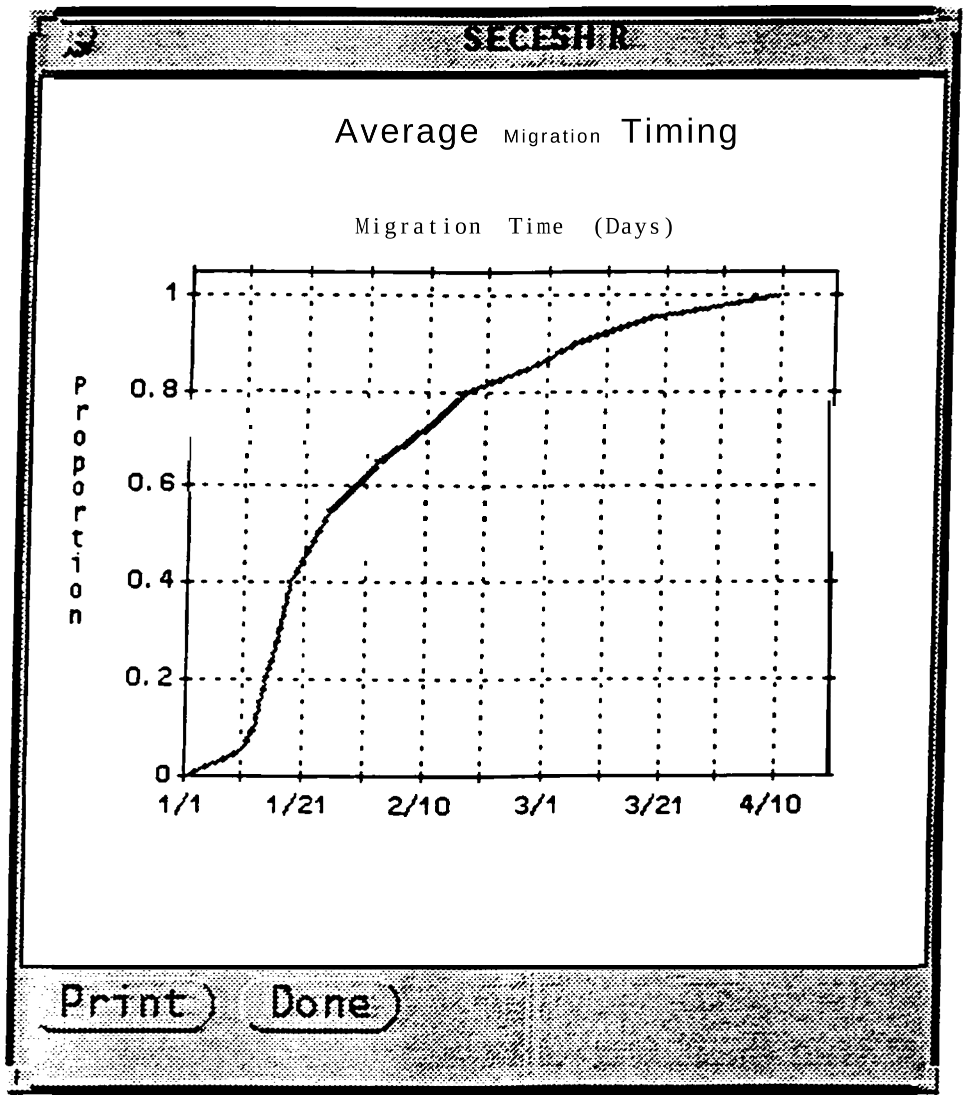
Cumulative detections normalized and then smoothed using a 5-point smoothing vs. detection dare.

Historical Data Plots



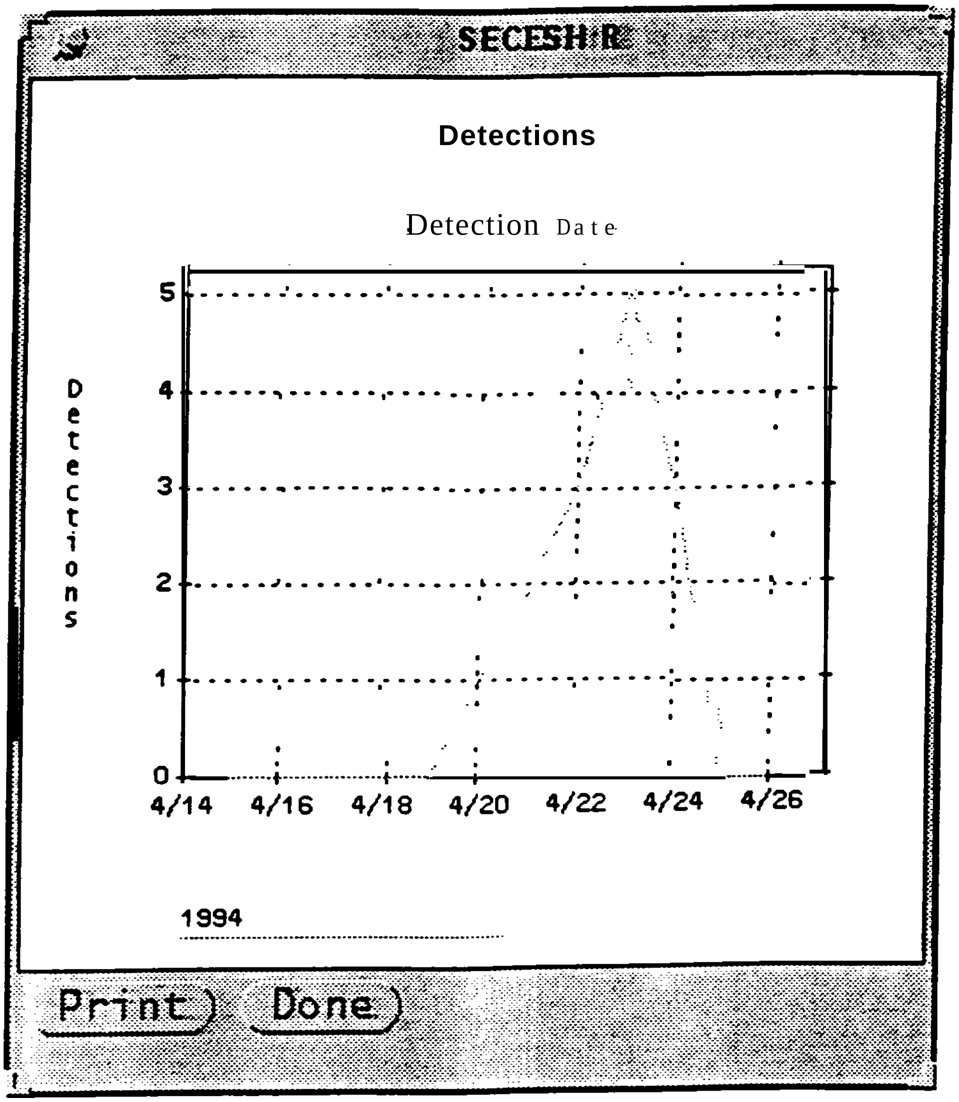
Slope (derivative) of cumulative **normalized** and smoothed **detections** vs. detection date. used in **LS** method.

Historical Data Plots



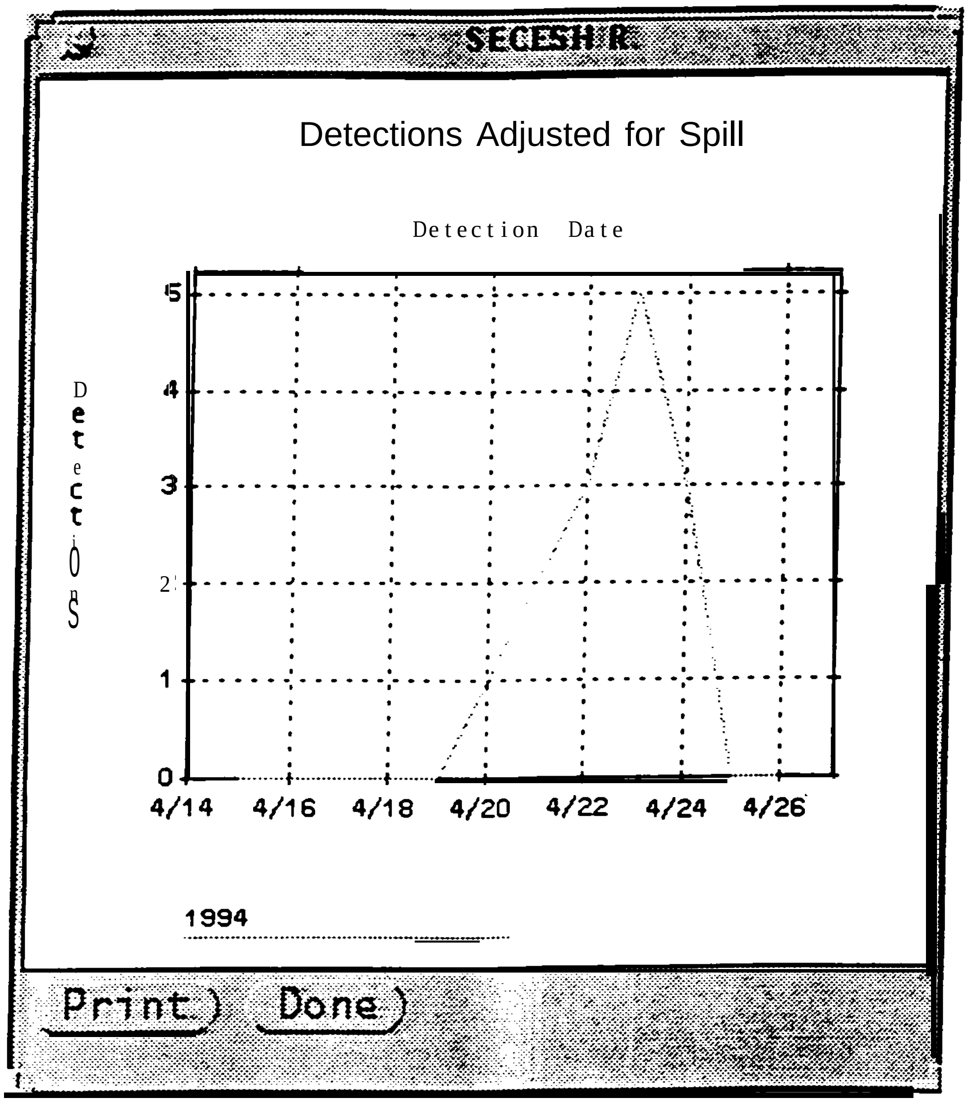
The average migration timing over historic years.

Current Data Plots



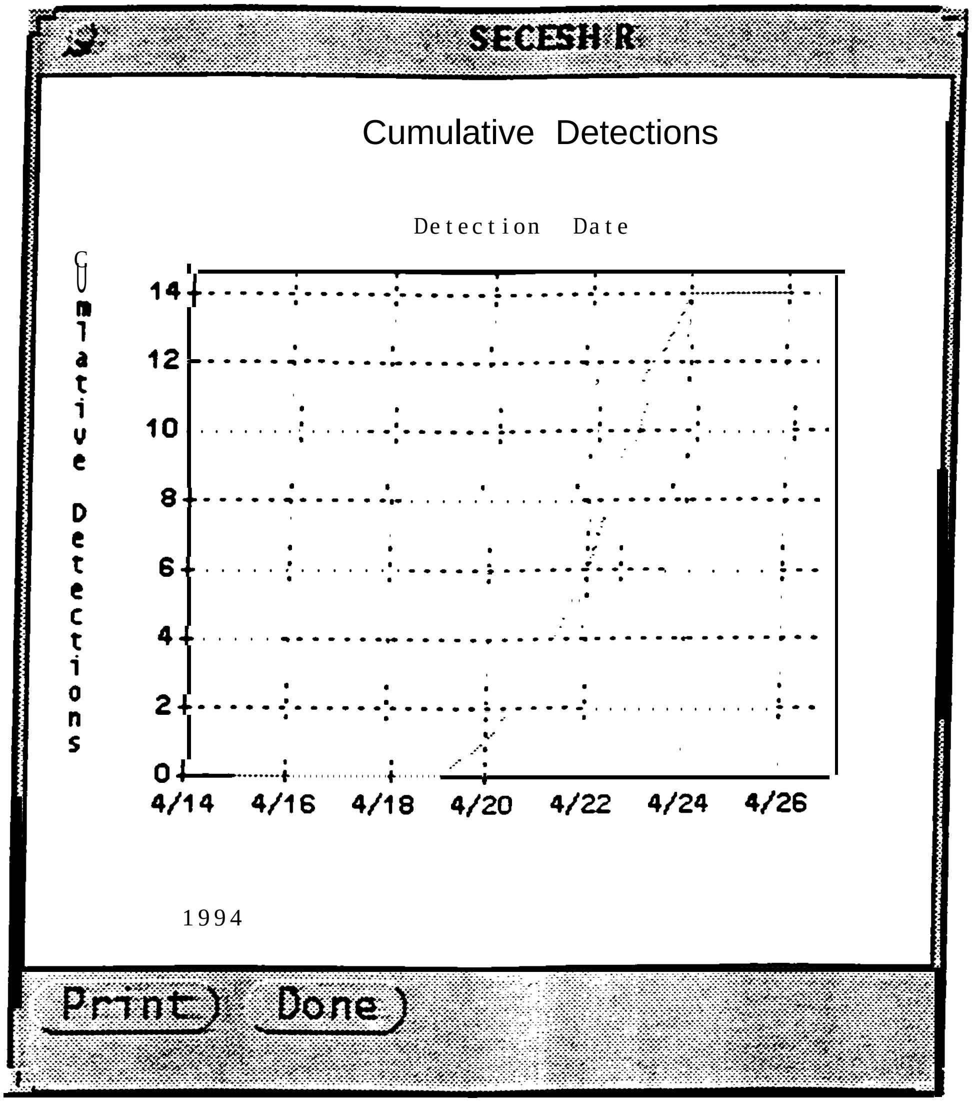
Detections to date vs. detection date for current year.

Current Data Plots



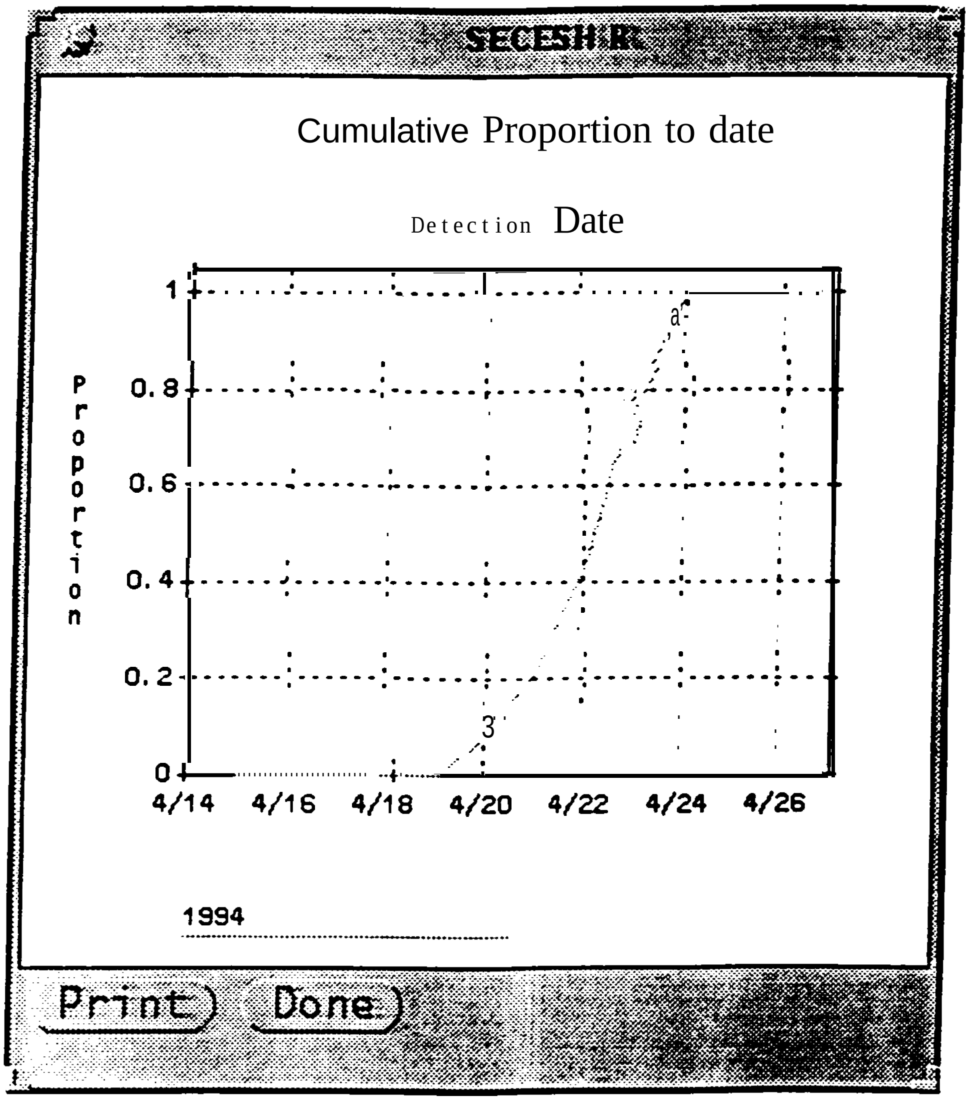
Detections to date adjusted for spill efficiency vs. detection date.

Current Data Plots



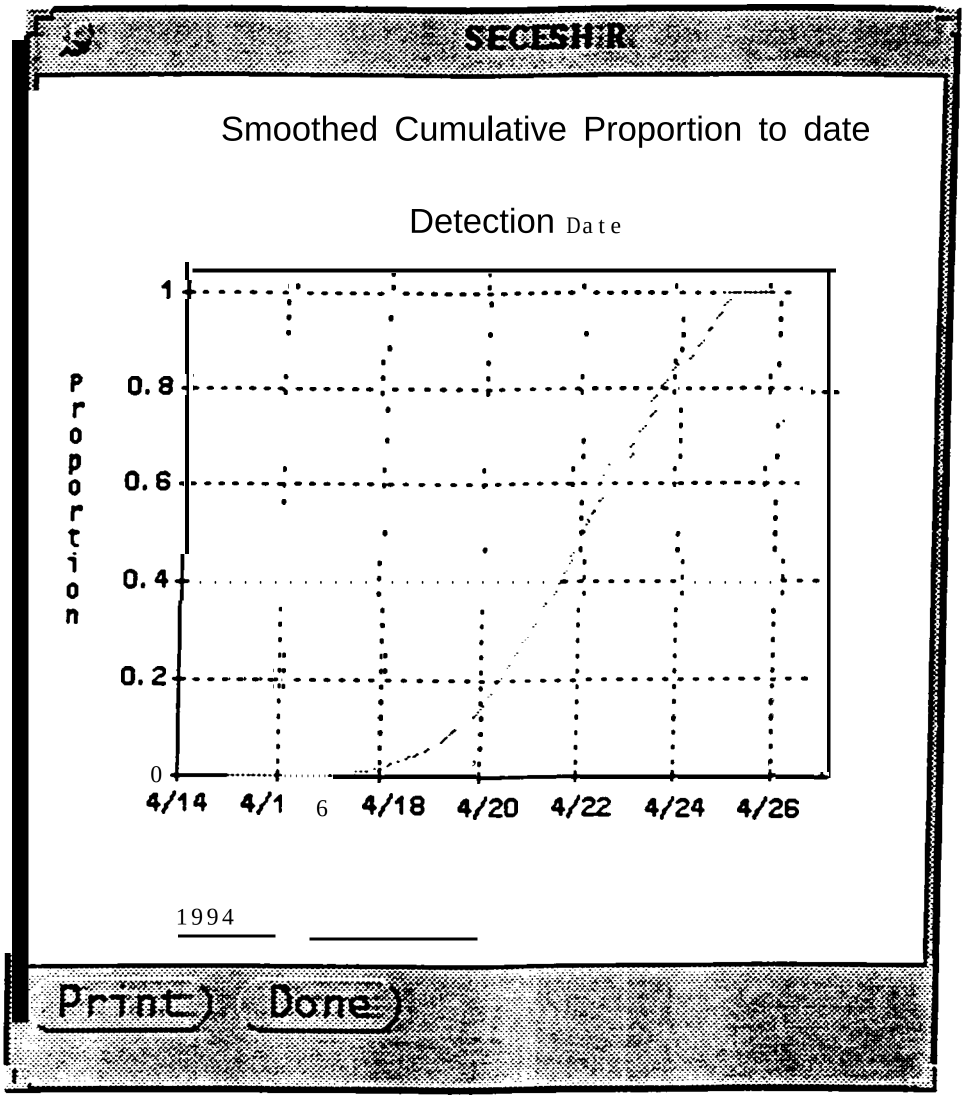
Cumulative detections (adjusted for spill) vs. detection date.

Current Data Plots



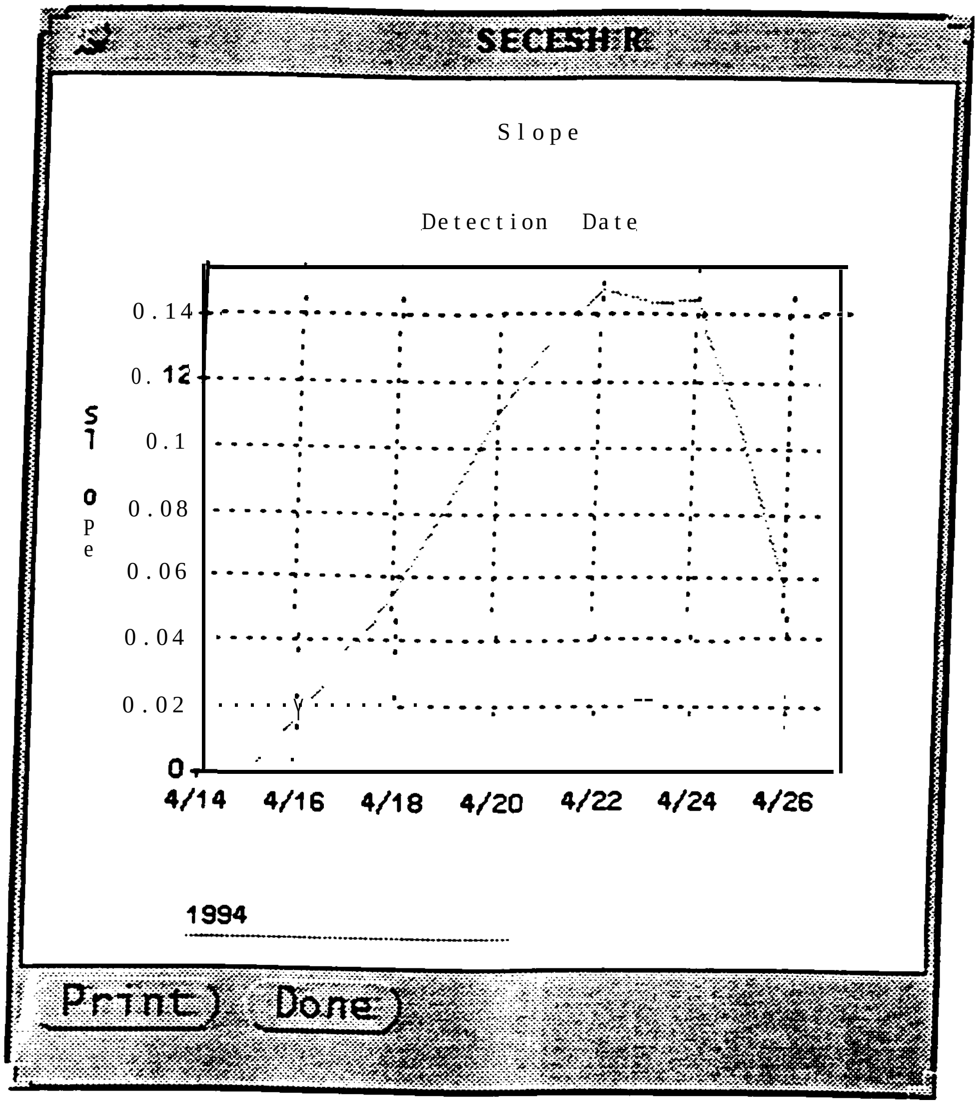
Cumulative detections normalized vs. detection date.

Current Data Plots



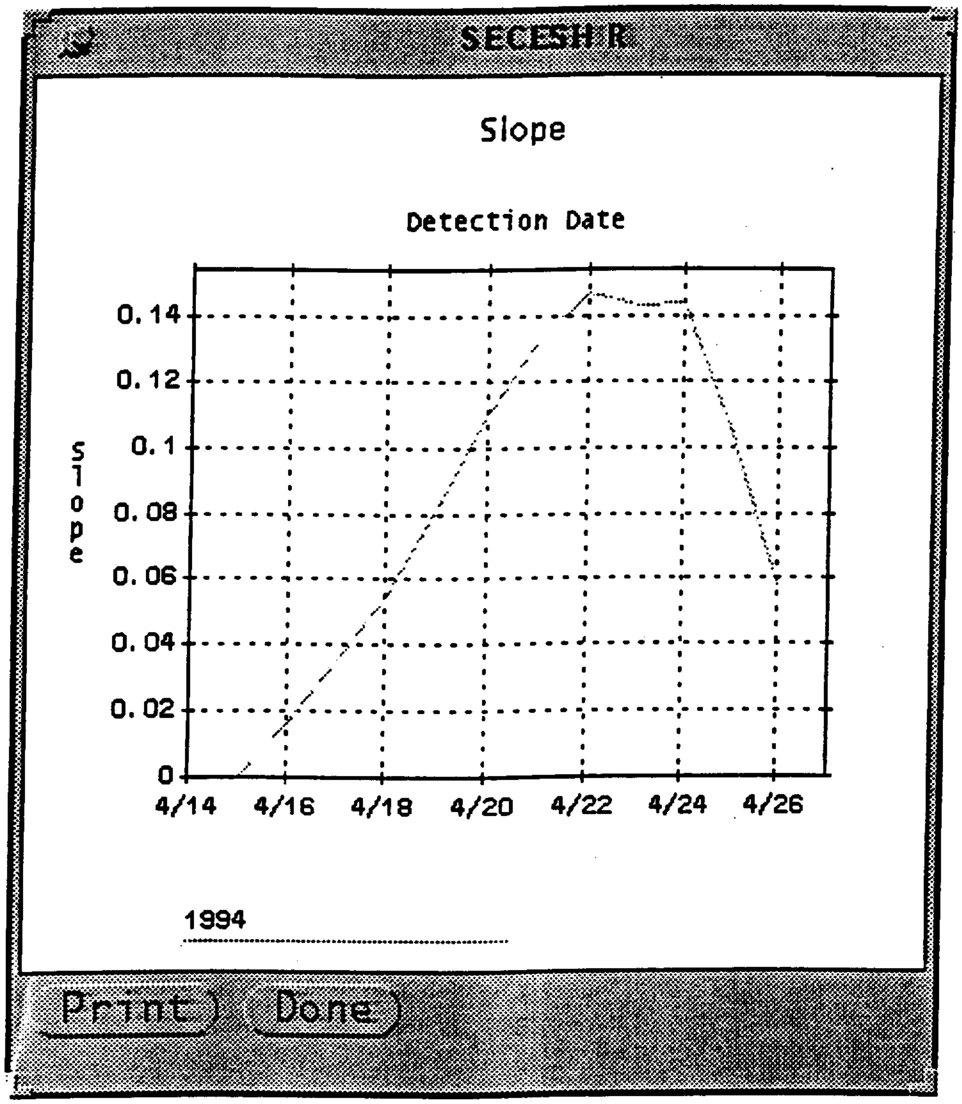
Cumulative detections smoothed using 5-point smoothing vs.detection daze.

Current Data Plots



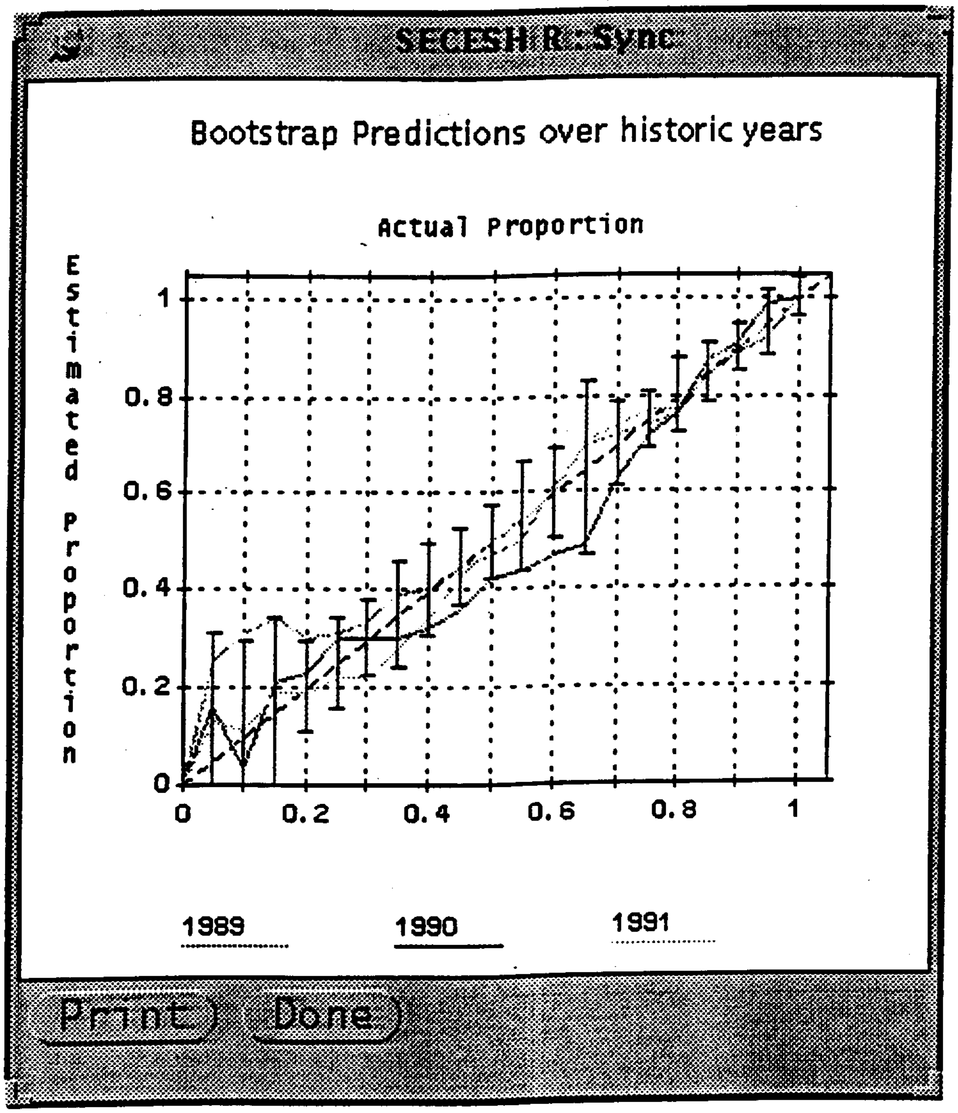
Slope (derivative) of cumulative normalized and smoothed detections vs. detection date. Used in "LS" method.

Current Data Plots



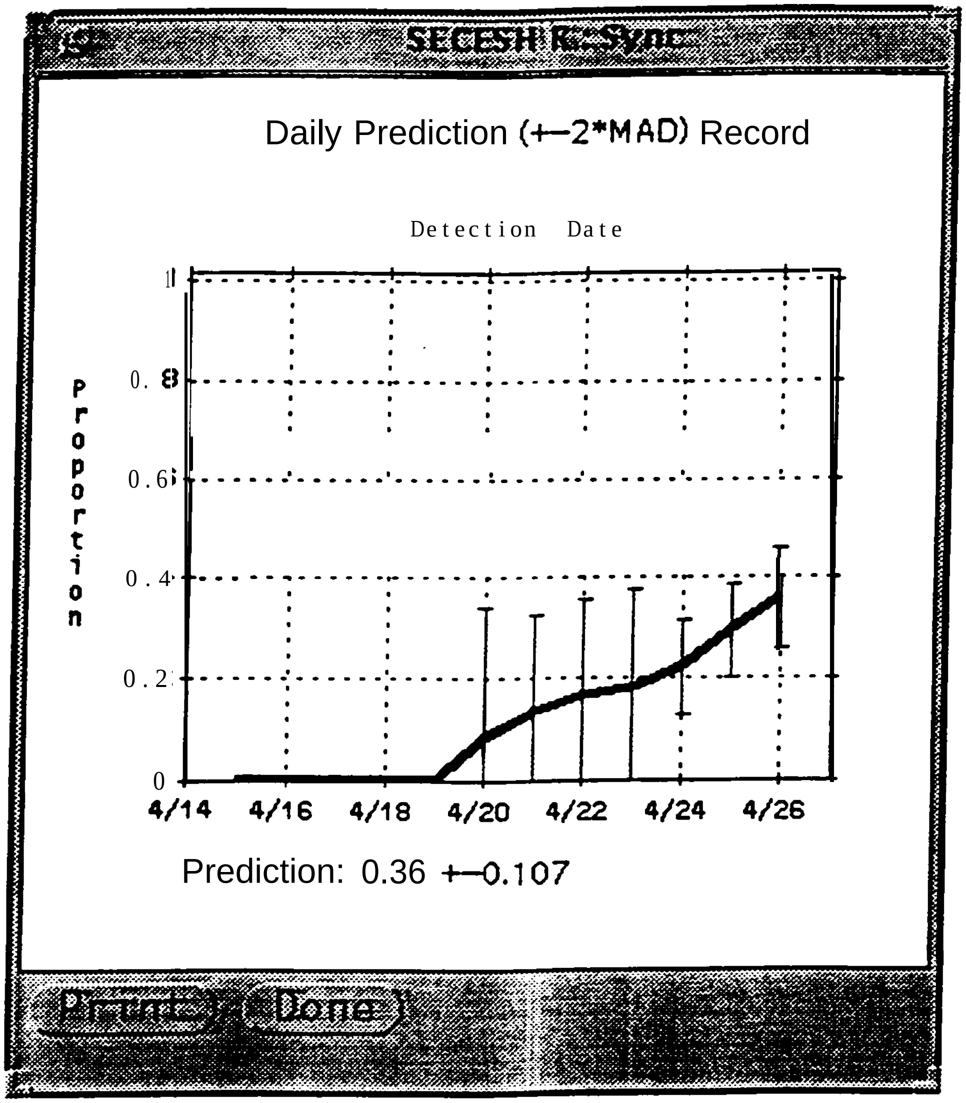
Slope (derivative) of cumulative normalized and smoothed detections vs. detection date. Used in "LS" method.

Results Plots



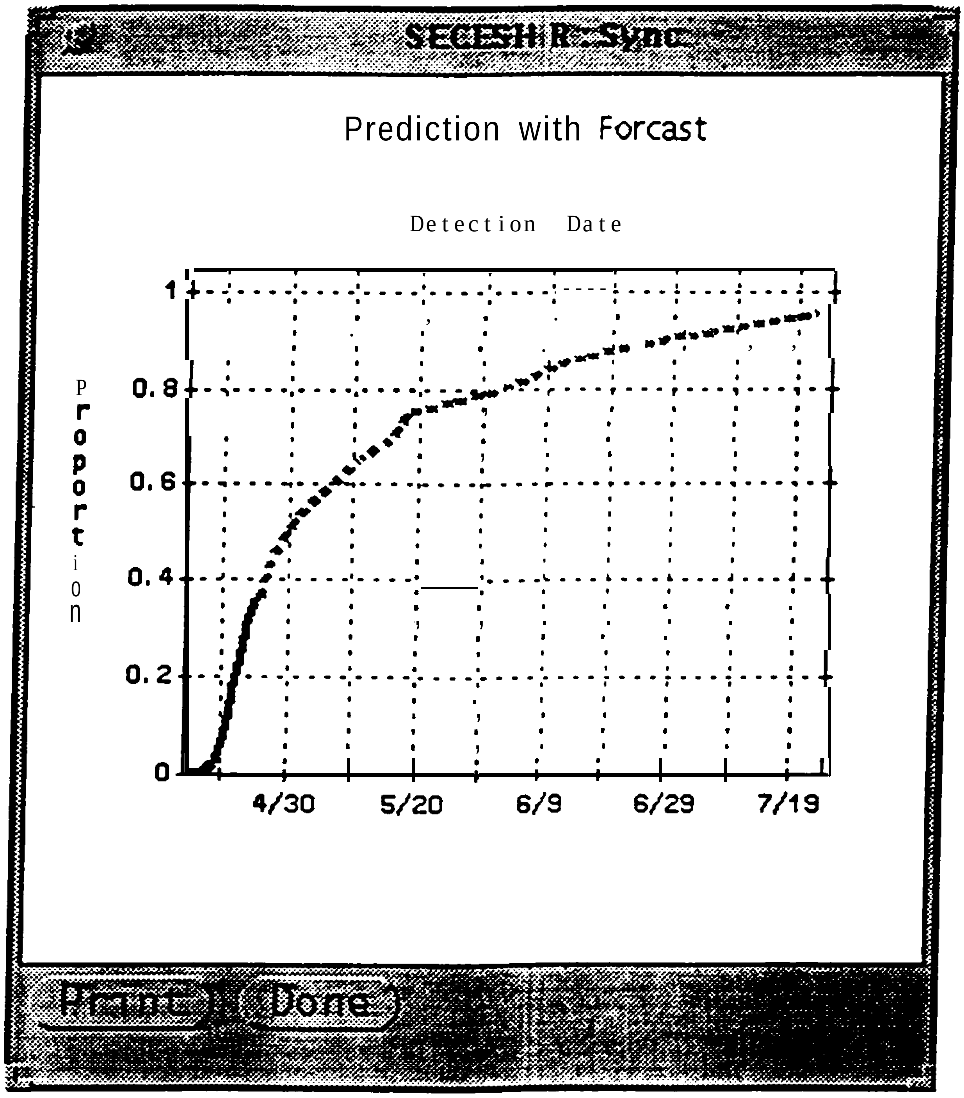
Bootstrapping predictions over historic years as a measure of the accuracy and precision of the prediction method.

Results Plots



The prediction for each day of the current run is plotted with error bars (plus or minus 2 times the Mean Absolute Deviation)

Results Plots



The current date fit to the predicted proportion in red plus the forecast based on historic migration timing averages in green.