

Simultaneous Forecasts of Multiple Interrelated Time Series with Markov Chain Model

Gongwei Chen, Ph.D., Washington State Caseload Forecast Council

ABSTRACT

In forecasting, there are often situations where several time series are interrelated: components of one time series can transition into and from other time series. A Markov chain forecast model may readily capture such intricacies through the estimation of a transition probability matrix, which enables a forecaster to forecast all the interrelated time series simultaneously. A Markov chain forecast model is flexible in accommodating various forecast assumptions and structures. Implementation of a Markov chain forecast model is straightforward using SAS[®]/IML. This paper demonstrates a real world application in forecasting community supervision caseload in Washington State. A Markov model was used to forecast five interrelated time series in the midst of turbulent caseload changes. This paper discusses the considerations and techniques in building a Markov chain forecast model at each step. Sample code using SAS[®]/IML is provided.

Anyone interested in adding another tool to their forecasting technique toolbox will find the Markov approach useful and have some unique advantages in certain settings.

INTRODUCTION

A discrete-time Markov chain is a stochastic process that consists of a finite number of states and transition probabilities among the different states. The process evolves through successive time periods. A first-order Markov chain process is characterized by the "Markov property", which states that the conditional probability distribution for the system at the next time period depends only on the current state of the system, and not on the state of the system at any previous time periods. Thus, such a Markov chain process is "memoryless". Consequently, it can be used for describing systems that follow a chain of linked events, where what happens next depends only on the current state of the system.

Markov chain models have wide applications in physics, chemistry, biology, medicine, economics, finance, and other disciplines. This paper presents a Markov chain model to forecast criminal justice caseload in Washington State, using SAS[®]/IML as the programming tool. The same methodology can be applied to many other areas of forecasting.

WASHINGTON STATE COMMUNITY SUPERVISION CASELOAD

Community Supervision is administered by the Washington State Department of Corrections (WADOC), and is an integral part of Washington's criminal justice system. The WADOC supervises certain offenders who have either been released from incarceration, or were sentenced directly to supervision in the community. A high percentage of offenders have conditions of supervision in their sentencing. These conditions are guided by public-safety considerations and engage each offender in programs to reduce risk of re-offending in the community.

At the Washington State Caseload Forecast Council, we are charged with forecasting the community supervision caseload, among many other program areas. These forecasts are generated independently to facilitate the budgeting process in Washington State. Washington State's community supervision program has undergone substantial changes in the past two decades. In response to such changes, we have adapted and utilized different models to forecasting this caseload. One of these models is a Markov chain model, which we started using in 2003.

At that time the community supervision caseload had three distinct supervision categories: Active, Inactive, and Monetary-Only. Active caseload includes those cases where the offenders are actively being supervised by the WADOC in the community. Inactive caseload includes those cases where offenders are no longer being actively supervised because of a court or board action, or because of

absconding. Monetary-Only caseload includes those cases where the offenders are supervised to pay restitution only.

Offenders can enter the supervision system at any supervision category, although the majority of them enter the active category first. Upon entering the Active caseload, they are evaluated for risk of future re-offense and are assigned to one of two levels: High Risk or Low Risk. While being evaluated for re-offense risk, an offender's risk level is briefly assigned as "Unclassified".

During supervision, the offenders' supervision category and risk level can change with good or bad behavior, new offense, illness, escape, and many other factors. As a result, any movement among the supervision categories and risk levels is possible: An active case may become an inactive case; a Monetary-Only case can re-offend and become an active case; a low-risk offender may be re-classified as high-risk, etc. Furthermore, cases may be terminated from any supervision category and risk level in any given month, and new cases are added to each category and risk level each month.

For budgeting and planning purposes, CFC was charged with forecasting not only all three supervision categories' caseload, but also the caseload under different risk levels separately (Unclassified, High Risk, Low Risk). So we must forecast a total of five components of the supervision caseload: High Risk Active (HRA), Low Risk Active (LRA), Unclassified Active (UNC), Monetary-Only (MON), and Inactive (INA).

From one month to another, the possible movements of the five components can be illustrated by the following chart. This process repeats into the future.

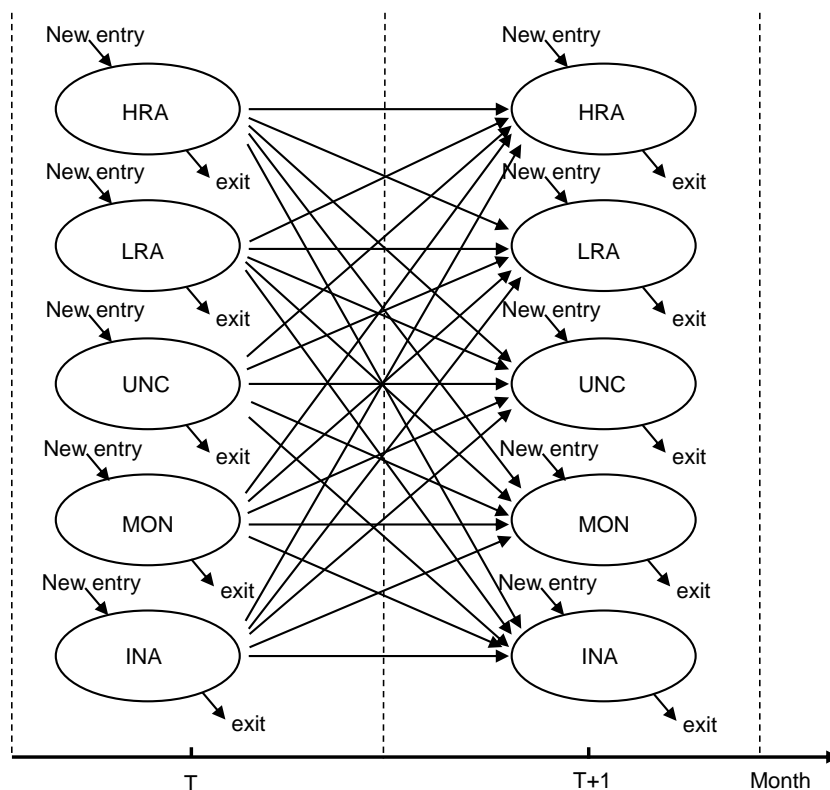


Figure 1. Supervision Caseload Movements

MARKOV MODEL DEVELOPMENT

The sequential flow of possible movements in Figure 1 appears to fit well into a Markov chain framework. Below are further considerations of suitability, and the steps to develop a Markov chain model to forecast the various components of the supervision caseload.

MARKOV CHAIN STATES

The five aforementioned components of the community supervision system are exhaustive and mutually exclusive. At any given time, an offender can only belong to one of the components; and the forgoing five components included all the supervision caseload WADOC was in charge of. In Markov chain model jargon, such exhaustive and non-overlapping components of the system are call “states”. A supervision offender can be in one, and only one, of these five states at any given time.

TRANSITION PROBABILITY MATRIX

In a first-order Markov chain process, a transition probability matrix exists that describes the probabilities of transitioning from one state to another in successive time periods. To facilitate discussion, the five states of the supervision caseload are denoted as follows:

State A: High Risk Active,

State B: Low Risk Active,

State C: Unclassified Active,

State D: Monetary-Only,

State E: Inactive.

The possibility of transitioning from one state to any other state of the system can then be described by a 5x5 matrix, where each row represents a state that an offender can move out of, and each column represents a state that an offender can move into. Given monthly data at the individual offender level, we can estimate such a matrix for each two consecutive months:

$$\begin{bmatrix} P_{AA} & P_{AB} & P_{AC} & P_{AD} & P_{AE} \\ P_{BA} & P_{BB} & P_{BC} & P_{BD} & P_{BE} \\ P_{CA} & P_{CB} & P_{CC} & P_{CD} & P_{CE} \\ P_{DA} & P_{DB} & P_{DC} & P_{DD} & P_{DE} \\ P_{EA} & P_{EB} & P_{EC} & P_{ED} & P_{EE} \end{bmatrix},$$

where P_{AA} is the probability of an offender in state A in month T to remain in the same state in month (T+1), while P_{AB} is the probability of an offender in state A in month T to transition to state B in month (T+1), and so on.

The sum of the first row in the matrix, $(P_{AA} + P_{AB} + P_{AC} + P_{AD} + P_{AE})$, is the total percentage of state A offenders at month T who remain in the community supervision caseload in one of the five states in month (T+1). Because some state A offenders are terminated from the caseload each month, $(P_{AA} + P_{AB} + P_{AC} + P_{AD} + P_{AE})$ is always less than 100%. This holds true for other rows of the matrix as well. In some other Markov processes though, the system is always in one of its states. Consequently each row of the transition probability matrix would add up to 100%.

We are able to estimate such a matrix between consecutive months because we have individual level of data: the data records the state a particular offender is in in all of the months he is under community supervision. Such individual level of data is crucial to estimate the transition probability matrix and to build a Markov model. If individual level of data is unavailable, there are numerical procedures that might enable one to obtain maximum likelihood or least square estimates of the transition probability matrix from aggregate data (see Dent and Ballitine 1971). Depending on these numerical procedures' performance, the lack of individual data might pose an obstacle in utilizing a Markov chain model.

Assuming that the transition probability matrix remains stable from month to month, and that in month T the number of offenders in the five states is A_T , B_T , C_T , D_T , and E_T , respectively, then the caseload in month (T+1) can be forecasted as follows:

$$[A_{T+1} \ B_{T+1} \ C_{T+1} \ D_{T+1} \ E_{T+1}] = [A_T \ B_T \ C_T \ D_T \ E_T] \cdot \begin{bmatrix} P_{AA} & P_{AB} & P_{AC} & P_{AD} & P_{AE} \\ P_{BA} & P_{BB} & P_{BC} & P_{BD} & P_{BE} \\ P_{CA} & P_{CB} & P_{CC} & P_{CD} & P_{CE} \\ P_{DA} & P_{DB} & P_{DC} & P_{DD} & P_{DE} \\ P_{EA} & P_{EB} & P_{EC} & P_{ED} & P_{EE} \end{bmatrix}, \quad (1)$$

where A_T is the number of state A offenders in month T, A_{T+1} is the number of state A offenders in month (T+1), etc.

For illustration purposes, the calculated monthly transition probabilities for the five states of supervision caseload during year 2004 are shown in Table 1 to Table 5, below. From these tables, one can average the transition probabilities for a certain time period, and use it in equation (1) to forecast future caseload, provided that one believes that the transition probabilities during the time period chosen are time-homogeneous and are likely to remain relatively stable into the future. Time homogeneity of transition probability matrices is discussed below.

	P_{AA}	P_{BA}	P_{CA}	P_{DA}	P_{EA}
Jan-04	91.9	1.28	7.46	0.12	2.05
Feb-04	93.2	0.88	6.75	0.18	1.65
Mar-04	92.0	1.30	8.96	0.20	2.41
Apr-04	92.4	1.14	9.22	0.17	2.15
May-04	91.0	1.02	8.12	0.16	1.90
Jun-04	92.5	1.12	10.97	0.35	2.69
Jul-04	92.2	0.75	11.99	0.23	2.26
Aug-04	92.4	0.83	13.90	0.39	2.24
Sep-04	92.0	0.93	11.06	0.28	2.31
Oct-04	91.5	0.87	10.56	0.23	2.31
Nov-04	92.4	0.94	12.74	0.51	2.41
Dec-04	91.8	0.75	10.08	0.46	2.35

Table 1. Transition into State A (%)

	P_{AB}	P_{BB}	P_{CB}	P_{DB}	P_{EB}
Jan-04	0.52	89.0	11.8	0.25	1.83
Feb-04	0.56	92.7	11.1	0.22	1.42
Mar-04	0.76	90.2	13.6	0.28	1.91
Apr-04	0.80	91.0	13.5	0.18	1.78
May-04	0.61	90.4	12.8	0.22	1.55
Jun-04	0.64	89.4	25.5	0.30	1.66
Jul-04	0.54	91.3	15.1	0.28	1.54
Aug-04	0.66	89.2	17.9	0.32	1.48
Sep-04	0.79	91.0	16.5	0.26	1.37
Oct-04	1.13	89.6	14.7	0.30	1.17
Nov-04	0.83	89.9	16.2	0.13	1.38
Dec-04	1.12	90.7	13.9	0.81	1.16

Table 2. Transition into State B (%)

	P_{AC}	P_{BC}	P_{CC}	P_{DC}	P_{EC}
Jan-04	0.00	0.00	64.8	0.06	0.64
Feb-04	0.00	0.00	71.5	0.07	0.60
Mar-04	0.00	0.00	66.2	0.04	0.43
Apr-04	0.00	0.00	64.7	0.05	0.44
May-04	0.00	0.00	59.9	0.03	0.55
Jun-04	0.00	0.00	50.2	0.07	0.53
Jul-04	0.00	0.00	59.5	0.03	0.52
Aug-04	0.01	1.26	55.7	0.00	0.42
Sep-04	0.01	0.00	61.1	0.05	0.57
Oct-04	0.01	0.00	61.6	0.00	0.41
Nov-04	0.00	0.00	58.0	0.13	0.51
Dec-04	0.01	0.01	63.2	0.12	0.36

Table 3. Transition into State C (%)

	P_{AD}	P_{BD}	P_{CD}	P_{DD}	P_{ED}
Jan-04	2.32	3.85	2.75	42.1	0.91
Feb-04	1.77	2.42	1.91	98.2	1.03
Mar-04	2.30	3.51	2.48	70.9	0.86
Apr-04	2.20	3.33	1.93	36.8	0.35
May-04	1.75	2.88	1.78	54.3	0.32
Jun-04	2.10	3.72	2.79	66.4	0.36
Jul-04	1.93	3.61	2.89	96.6	0.31
Aug-04	0.89	1.24	1.22	82.9	0.09
Sep-04	0.61	0.75	0.77	56.5	0.06
Oct-04	0.37	0.57	0.50	53.2	0.08
Nov-04	0.34	0.41	0.76	45.5	0.06
Dec-04	0.27	0.25	0.12	57.1	0.06

Table 4. Transition into State D (%)

	P _{AE}	P _{BE}	P _{CE}	P _{DE}	P _{EE}
Jan-04	2.70	2.33	5.70	0.01	92.5
Feb-04	2.55	1.79	5.76	0.01	93.9
Mar-04	2.62	1.75	4.81	0.02	92.9
Apr-04	2.46	1.97	4.65	0.03	94.0
May-04	4.29	3.06	13.65	0.03	94.5
Jun-04	2.41	1.94	6.13	0.02	93.5
Jul-04	3.03	1.85	8.86	0.00	94.4
Aug-04	2.74	1.88	6.11	0.11	93.1
Sep-04	2.65	1.75	5.28	0.00	94.8
Oct-04	3.28	2.13	7.06	0.04	95.1
Nov-04	2.65	1.77	6.45	0.00	93.3
Dec-04	3.05	2.11	7.62	0.00	94.3

Table 5. Transition into State E (%)

TIME-HOMOGENEITY OF TRANSITION PROBABILITY MATRICES

With N months of historical data, we can estimate (N-1) transition probability matrices. These matrices are rarely identical though. Statistical procedures exist to test whether the transition probability matrices are statistically different from one period to another (see Basawa and Rao 1980). If the transition probability matrices are found not statistically different from one period to another, time-homogeneity is established for the transition probability matrices. Should this be the case, the average of the (N-1) estimated transition probability matrices is the most efficient estimator for the transition probability matrix during the N months sampled.

In practice, domain knowledge often plays an important role in determining the time-homogeneity of transition probability matrices. In forecasting Washington State's supervision caseload, legislative and policy insights are crucial. Usually some attempts will be made each year, during the legislative session, to amend existing law or introduce new law that affects Washington State's criminal justice system. Furthermore, WADOC has some discretion over its supervision policies and practices. Legislative/policy changes can change the dynamics of the caseload and affect the transition probabilities. Once a legislative or policy change has occurred, and substantial impact on the caseload is expected, one needs to closely monitor the evolution of the transition matrices and determine if there has been a paradigm change and all previous transition probability matrices under the old law/policy need to be discounted, and only latest data should be used in estimating the transition probability matrices.

One also needs to have intimate knowledge of the data one works with. All transition probabilities that seem out of the ordinary need to be investigated for causes. For instance, if one notices the transition probability matrix of a particular month is statistically different from an average month, one should not be tempted to automatically throw it out based on statistical grounds alone. If further investigation shows the cause of the abnormal month is underreporting of data, and the underreporting was subsequently caught up in the ensuing months, then one's decision might be to keep this "outlying" transition probability matrix when calculating an average of transition probability matrices.

So, in-depth domain knowledge of the area one forecasts and intimate working knowledge of one's data can be just as important as statistical tests in determining if time-homogeneity is an appropriate assumption for a Markov chain model.

We have individual level of data for WADOC supervision caseload starting in September 2002. Figure 2 shows the historical supervision caseload by its five states, as of September 2004, when we were on schedule to produce a forecast in October 2004.

A visual examination of the historical data indicates that transition probability matrices during the period of September 2002 to September 2004 are probably not homogeneous, due to the unusual caseload movements during the following two periods:

1. May 2003 to July 2003: state B (Low Risk Active), state D (Monetary-Only), and state E (Inactive) caseload had all experienced some uncharacteristic changes.
2. December 2003 to May 2004: The precipitous drop of the state D (Monetary-Only) caseload in early 2004 was caused by a legislative change. A bill was passed in the Washington State Legislature in 2003, which stipulated that WADOC shall transfer all Monetary-Only cases to the county clerk's offices in the state. The transfer began in January 2004 and by September 2004 the bulk of cases had been transferred. By consulting WADOC experts, we learned we should expect all Monetary-Only cases to be transferred by January 2005.

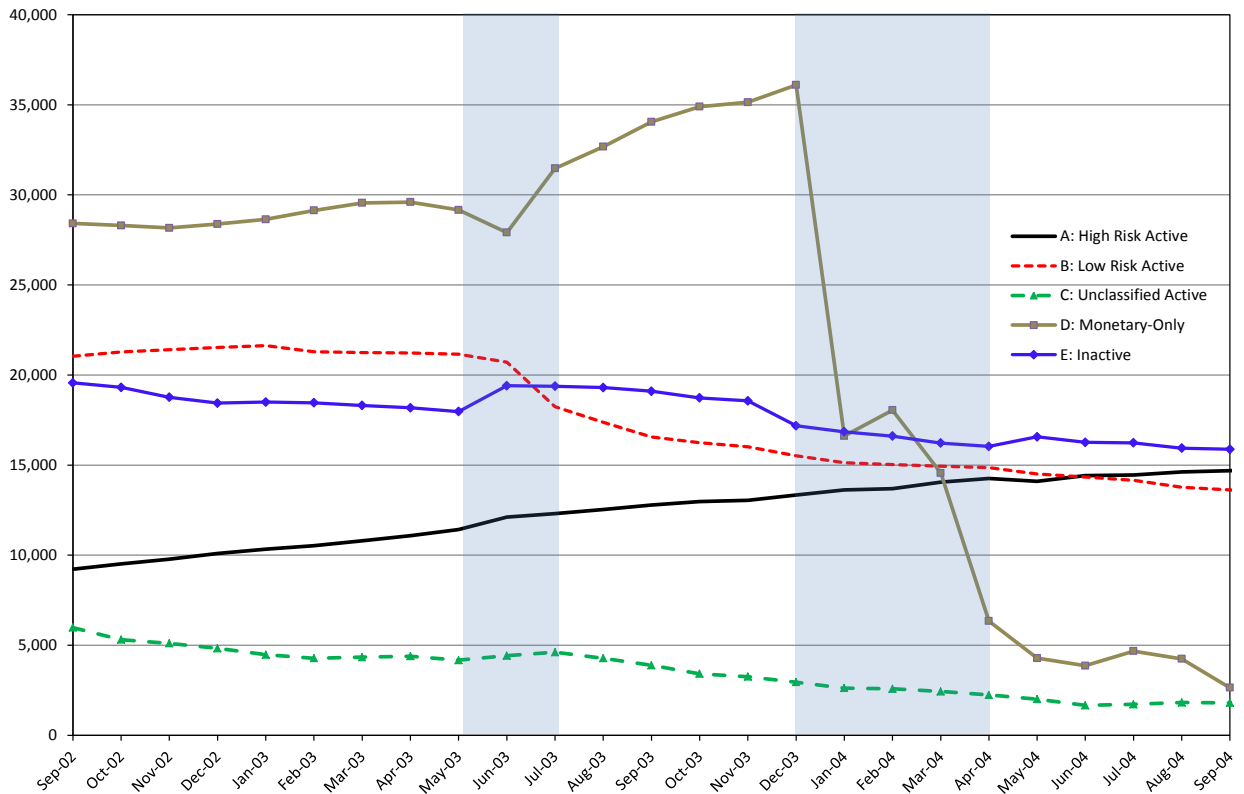


Figure 2. Historical Supervision Caseload by Markov Chain States

Considering that transition probability matrices probably have been affected significantly by the 2003 legislative change, it is prudent to use relatively recent data only and discard all data points prior to May 2004 for the purpose of estimating transition probability matrices to forecast future caseloads. This approach is based on domain knowledge, rather than statistical tests. It leaves only five usable data points (May to September 2004), and a total of only four transition probability matrices can be estimated. But such lack of longer history and usable data points is a predicament that is often faced by practitioners. With only four transition probability matrices, it is impossible to perform any meaningful statistical test regarding the homogeneity of these matrices. Homogeneity is assumed though, based on domain knowledge.

By averaging the transition probabilities (June to September 2004) given in Table 1 to Table 5, we calculate the following transition probability matrix (values are percentages) to be used for forecasting future caseloads¹.

¹ In Table 4, value of P_{DD} for July 2004 appears to be an outlier, so it is removed from the data in calculating average transition probability here.

$$\begin{bmatrix} P_{AA} & P_{AB} & P_{AC} & P_{AD} & P_{AE} \\ P_{BA} & P_{BB} & P_{BC} & P_{BD} & P_{BE} \\ P_{CA} & P_{CB} & P_{CC} & P_{CD} & P_{CE} \\ P_{DA} & P_{DB} & P_{DC} & P_{DD} & P_{DE} \\ P_{EA} & P_{EB} & P_{EC} & P_{ED} & P_{EE} \end{bmatrix} = \begin{bmatrix} 92.251 & 0.658 & 0.003 & 1.382 & 2.708 \\ 0.905 & 90.227 & 0.314 & 2.330 & 1.855 \\ 11.979 & 18.739 & 56.640 & 1.919 & 6.594 \\ 0.313 & 0.292 & 0.036 & 68.564 & 0.033 \\ 2.375 & 1.511 & 0.509 & 0.206 & 93.938 \end{bmatrix}$$

One complication to the forecast model is that we expect that State D will cease to exist starting in January 2005: all Monetary-Only caseload would be transferred out of WADOC. Fortunately, it is straightforward to handle the termination of one or more states in a Markov model: The transition probability matrix above can be used for forecasting the remaining three months of 2004. To forecast caseload in January 2005 and beyond, we just need to reset the forecast of state D caseload to zero in each and every month. With state D caseload set to zero in each month, the fourth row of the matrix becomes irrelevant, because the product of zero and any probability is still zero; consequently no state D case is transferred to other states in the forecasting of next period's caseload.

Resetting the forecast of state D caseload to zero in January 2005 and beyond also has the following effect on the forecasts: The cases transferred out of state A, B, C, E and into D would be treated as being terminated from state A, B, C, and E, respectively, and zero caseload is added to state D.

ENTRIES AND EXITS

In addition to transitions among the five states, there are new offenders entering into, and existing offenders exiting out of, each state during any given month. No special consideration is needed for modeling exits. As explained before, each row of the transition probability matrix is usually than 100%. The difference between the sum of probabilities in a certain row and 100% is the percentage of caseload terminated from that corresponding state during a month. When the transition probability matrix is used in equation (1), exits are automatically removed from the next time period's caseloads.

However, new entries into each state need to be separately modeled. Depending on the specific situations where a Markov chain model is applied, various methods can be used to forecast future entries. In the case of WADOC supervision caseload, we've observed that a strong predictor of monthly admissions is the number of working days in a particular month. Hence, we use linear regression models with the number of working days in a particular month as the independent variable to forecast entries into each of the five states by month. Since the method of forecasting entries is not unique to modeling a Markov chain process, this paper does not delve into further details of forecasting entries.

IMPLEMENTATION USING SAS®/IML

SAS®/IML provides a user-friendly tool to implement a Markov chain model. The sample code below illustrates the steps from reading in inputs (last actual, transition probability matrix, and new entries), to forecasting sequentially into the future.

```

DATA BasePopulation; /* 09/2004 data is last actual */
format date Monyy5.;
input
date Monyy5.  A      B      C      D      E;      cards;
Sep04        14683 13627 1800  2640  15878
;RUN;

DATA NewEntry (drop=date);
input
date Monyy5.  A      B      C      D      E;      cards;
Oct04         532   441   671   20   30
Nov04         488   416   553   15   34
Dec04         572   399   625   10   33
Jan05         500   394   547    0   34
Feb05         507   391   547    0   34
Mar05         621   470   662    0   41

```

Apr05	574	427	605	0	37
May05	581	424	605	0	37
Jun05	616	441	633	0	39
Jul05	567	398	576	0	35
Aug05	660	455	662	0	41
Sep05	609	412	605	0	37
Oct05	617	410	605	0	37
Nov05	564	368	547	0	34
Dec05	631	404	605	0	37
Jan06	607	382	576	0	35
Feb06	583	360	547	0	34
Mar06	714	433	662	0	41
Apr06	627	374	576	0	35
May06	698	408	633	0	39
Jun06	705	405	633	0	39
Jul06	648	366	576	0	35
Aug06	752	417	662	0	41
Sep06	661	360	576	0	35
Oct06	734	393	633	0	39
Nov06	641	337	547	0	34
Dec06	681	352	576	0	35

;RUN;

DATA TransitionMatrix;

input

A B C D E; cards;

```
92.251  0.658  0.003  1.382  2.708
0.905   90.23  0.314  2.330  1.855
11.979  18.74  56.64  1.919  6.594
0.313   0.292  0.036  66.38  0.033
2.375   1.511  0.509  0.206  93.938
```

;RUN;

PROC IML;

```
use        NewEntry;                read all into    NewEntry;
use        TransitionMatrix;        read all into    TransitionMatrix;
use        BasePolulation;          read all into    BasePolulation;
```

BasePolulation=BasePolulation[1, 2:6];

NewEntry=NewEntry[,1:5];

```
fc410=BasePolulation*TransitionMatrix/100+NewEntry[1, ];
fc411=fc410*TransitionMatrix/100+NewEntry[2, ];
fc412=fc411*TransitionMatrix/100+NewEntry[3, ];
fc501=fc412*TransitionMatrix/100+NewEntry[4, ]; fc501[1,4]=0;
fc502=fc501*TransitionMatrix/100+NewEntry[5, ]; fc502[1,4]=0;
fc503=fc502*TransitionMatrix/100+NewEntry[6, ]; fc503[1,4]=0;
fc504=fc503*TransitionMatrix/100+NewEntry[7, ]; fc504[1,4]=0;
fc505=fc504*TransitionMatrix/100+NewEntry[8, ]; fc505[1,4]=0;
fc506=fc505*TransitionMatrix/100+NewEntry[9, ]; fc506[1,4]=0;
fc507=fc506*TransitionMatrix/100+NewEntry[10,]; fc507[1,4]=0;
fc508=fc507*TransitionMatrix/100+NewEntry[11,]; fc508[1,4]=0;
fc509=fc508*TransitionMatrix/100+NewEntry[12,]; fc509[1,4]=0;
fc510=fc509*TransitionMatrix/100+NewEntry[13,]; fc510[1,4]=0;
fc511=fc510*TransitionMatrix/100+NewEntry[14,]; fc511[1,4]=0;
fc512=fc511*TransitionMatrix/100+NewEntry[15,]; fc512[1,4]=0;
```



```

fc601=fc512*TransitionMatrix/100+NewEntry[16,]; fc601[1,4]=0;
fc602=fc601*TransitionMatrix/100+NewEntry[17,]; fc602[1,4]=0;
fc603=fc602*TransitionMatrix/100+NewEntry[18,]; fc603[1,4]=0;
fc604=fc603*TransitionMatrix/100+NewEntry[19,]; fc604[1,4]=0;
fc605=fc604*TransitionMatrix/100+NewEntry[20,]; fc605[1,4]=0;
fc606=fc605*TransitionMatrix/100+NewEntry[21,]; fc606[1,4]=0;
fc607=fc606*TransitionMatrix/100+NewEntry[22,]; fc607[1,4]=0;
fc608=fc607*TransitionMatrix/100+NewEntry[23,]; fc608[1,4]=0;
fc609=fc608*TransitionMatrix/100+NewEntry[24,]; fc609[1,4]=0;
fc610=fc609*TransitionMatrix/100+NewEntry[25,]; fc610[1,4]=0;
fc611=fc610*TransitionMatrix/100+NewEntry[26,]; fc611[1,4]=0;
fc612=fc611*TransitionMatrix/100+NewEntry[27,]; fc612[1,4]=0;

```

```

Forecast=fc410//fc411//fc412//fc501//fc502//fc503//fc504//fc505//fc506//f
c507//fc508//fc509//fc510//fc511//fc512//fc601//fc602//fc603//fc604//fc60
5//fc606//fc607//fc608//fc609//fc610//fc611//fc612;

```

```

create Forecast from Forecast;      append from Forecast;
quit;

```

The forecast generated is plotted in Figure 3. Since state D caseload was soon to be terminated altogether, it is not plotted. The results of the Markov chain forecast model is that state A offenders are expected to continue its upward trajectory, state B and E offenders are expected to continue to decline, while state C offenders are projected to remain relatively stable.

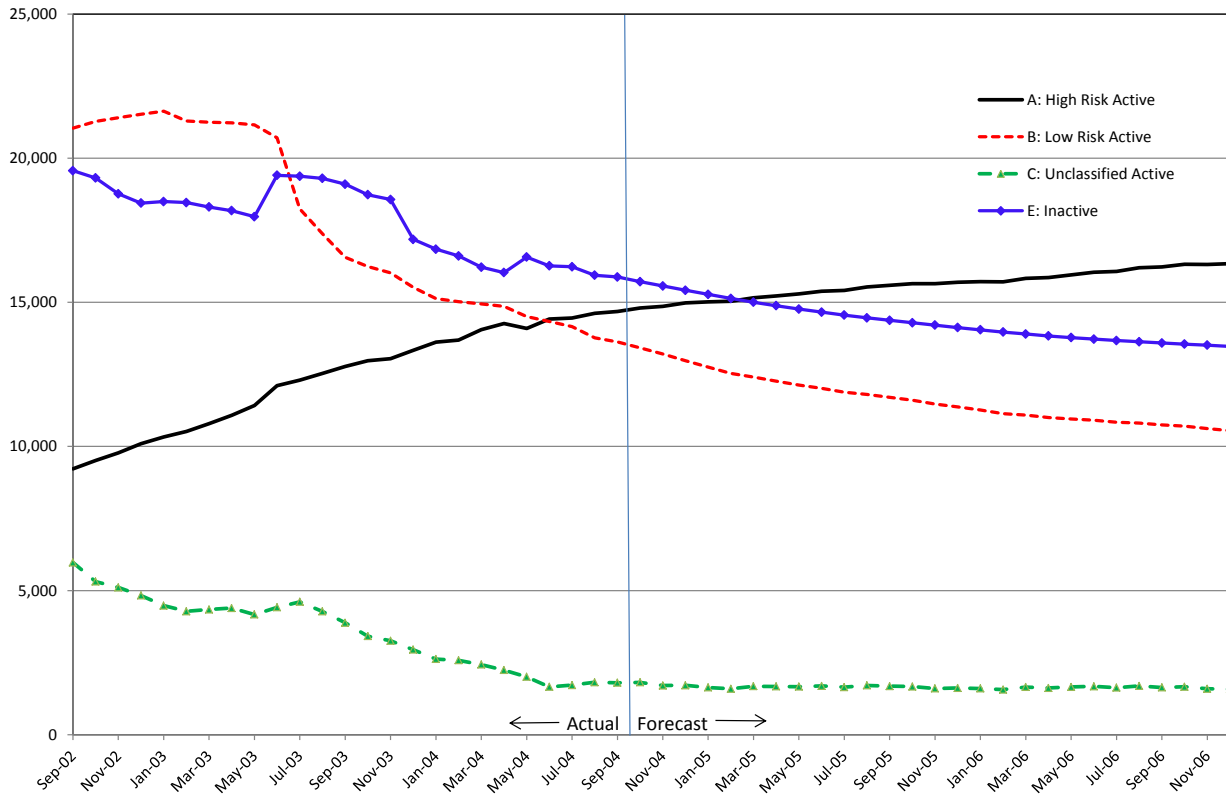


Figure 3. Markov Chain Model Forecast for State A, B, C, E

SENSITIVITY ANALYSIS

The transition probability matrix assumption and the new entries assumption jointly determine the outcome of the forecasts. If one has competing assumptions for new entries, or one is not certain how many data points to go back to estimate transition probability matrices, the author highly recommend that one conduct a sensitivity analysis to see how different assumptions change the forecasts. One needs to be alerted if the outcome of the Markov model is highly sensitive to small changes in the inputs.

For illustration purposes, we can use an alternative transition probability matrix estimated using data from January to September 2004 as an alternative assumption to see how different the forecasts would be from the original forecasts, whose transition probability matrix is estimated using data from June to September 2004. The alternative transition probability matrix is as follows (values are percentages):

$$\begin{bmatrix} P_{AA} & P_{AB} & P_{AC} & P_{AD} & P_{AE} \\ P_{BA} & P_{BB} & P_{BC} & P_{BD} & P_{BE} \\ P_{CA} & P_{CB} & P_{CC} & P_{CD} & P_{CE} \\ P_{DA} & P_{DB} & P_{DC} & P_{DD} & P_{DE} \\ P_{EA} & P_{EB} & P_{EC} & P_{ED} & P_{EE} \end{bmatrix} = \begin{bmatrix} 92.162 & 0.654 & 0.002 & 1.764 & 2.828 \\ 1.026 & 90.474 & 0.140 & 2.813 & 2.035 \\ 9.825 & 15.303 & 61.513 & 2.059 & 6.772 \\ 0.232 & 0.258 & 0.044 & 67.170 & 0.025 \\ 2.184 & 1.614 & 0.525 & 0.476 & 93.734 \end{bmatrix}$$

A comparison of the forecasts using the original and the alternative assumption is shown in Figure 4. It shows that, keeping the new entry assumptions unchanged, the original and alternative transition probability matrices produce relatively similar forecasts for state A, B, C, E offenders. The biggest difference is in the forecast of state A offenders. The alternative forecast is about 3% lower than the original at the end of the forecast horizon. The result of this exercise is reassuring because although the state D caseload underwent major changes during January 2004 to September 2004, and one may reasonably suspect the forecasts generated by the two different transition probability matrices might differ considerably, it turns out that changes in state D caseload does not appear to unduly affect the forecasts of caseloads in state A, B, C, and E.

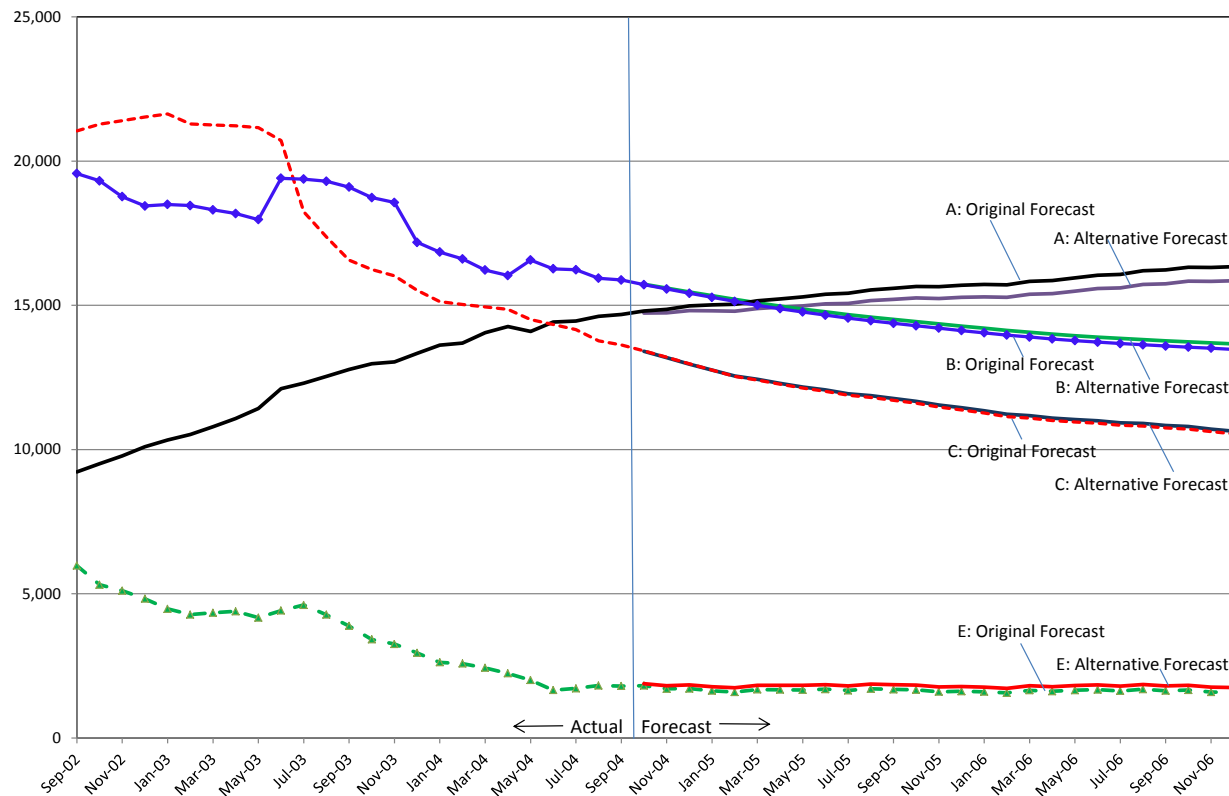


Figure 4. Sensitivity Analysis

MODEL REFINEMENT AND RELAXATION OF ASSUMPTIONS

There are various actions one can take to fine-tune the Markov chain model or relax its assumptions. A relatively easy one, when there are potentially many states in a Markov model, is to experiment with differentiating and combining different states to either generate forecasts with more granularities or achieve less model complexity.

Another area of model refinement, albeit more involved, is to relax the time-homogeneity assumption for the transition probability matrices in future time periods. Linear, or, other more sophisticated models can be built to forecast individual elements of future transition probability matrices, so that these transition probability matrices are time-dependent. One scenario such added complexity might be warranted is when one has several years of data, and historical data shows strong seasonality.

However, from a practitioner's perspective, one of the appeals of the Markov model is its simplicity and robustness, extra bells and whistles should be evaluated to see if they warrant the added complexity.

DISCUSSION/CONCLUSION

This paper introduces the Markov chain forecast model as a useful tool for situations where multiple time series are interrelated. Through the estimation of a transition probability matrix, a Markov chain forecast model may readily capture the transitions among the various states, and enable a forecaster to forecast all the interrelated time series simultaneously.

The case study of WADOC's supervision caseload illustrates that in a Markov chain forecast model, just like in other types of forecast models, domain knowledge and understanding of underlying data are as important as the mathematical/statistical side of a model building process. The case study also demonstrates that a Markov chain forecast model can be very flexible in accommodating various forecast assumptions and structures.

Another desirable feature of a Markov chain forecast model is its ability to build a model where historical data is limited. In the case study, we estimated a transition probability matrix using just five data points. Despite the turbulent changes of caseload in one of the states, the sensitivity analysis shows that the forecasts generated for caseloads in other states are not unduly influenced. The ability of building a model with limited data points can be desirable in forecasting areas where external shock to the system is frequent and accumulating enough data points to build an ARIMA type of model is a luxury.

A Markov chain forecast model can be easily implemented using SAS[®]/IML. For the right situation, it is a handy tool in a forecaster's toolbox.

REFERENCES

1. Dent, W., and Ballitine, R. (1971) A review of the estimation of transition probabilities in Markov chain. *Australian Journal of Agricultural Economics*, 15: 69–81.
2. Basawa, I., and Rao, B. *Statistical Inference for Stochastic Process*. London: Academic Press, 1980.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Gongwei Chen
Enterprise: Washington State Caseload Forecast Council
E-mail: gongwei.chen@cfc.wa.gov

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.