Original Research Article

Selection of Optimum Number of States for a Hidden Markov Manpower Model in a Departmentalized Framework

Abstract

Hidden Markov models are powerful tools used in studying various dynamic systems, especially where internal transitions are not directly observable. In statistical manpower planning, these models help capture the effects of hidden heterogeneity on personnel transitions across states of the system. However, in practice, a significant challenge is determining the number of hidden states within the model from data, as this choice is often made subjectively because observable data on hidden states are nonexistent. This study presents systematic approaches to handling this problem, in a general departmentalized framework, through two search procedures. The proposed search procedures are formulated to allow any suitable statistical tools. Specifically, Likelihood Ratio statistic, Akaike Information Criterion and Bayesian Information Criterion were applied through the procedures to identify the most suitable number of hidden states for three Hidden Markov manpower model applications reviewed. All the procedures with the various statistical tools employed gave the same and consistent results. In particular, the optimum number of hidden states was found to be two for all the manpower datasets analyzed, raising curiosity about the existence of general threshold points beyond which the addition of more hidden states in hidden Markov manpower models has no significant gain. The procedures are adaptable to other application areas, outside manpower systems, where application of hidden Markov models poses the same problem of choice of number of hidden states.

Commented [HP1]: Can you please update as “to determine”?

Keywords: Hidden Markov model, manpower data, EM algorithm, search procedure, optimum number of state, departmentalized manpower

Commented [HP2]: Should it be “states”?

1. Introduction

Hidden Markov manpower model (HMMM) is an essential statistical tool for understanding complex workforce systems. Human resource or workforce management involves the strategic allocation of people and resources, tracking attendance and ensuring compliance with workplace laws and regulations. Ultimately, the objectives are to optimize productivity and reduce risk, and, as well, improve employee satisfaction. Statistical manpower planning has been carried out over many decades now, through modelling and control, to achieve these objectives. In the course of this, many factors inherent in manpower systems have been taken into consideration. One such factor is workforce heterogeneity. Ugwuowo and McClean [1], Guerry [2] and Udom and Ebedoro [3] are among researchers who identified that manpower systems can have two parts: observable and hidden parts. The observable part is the part that can be observed through available data. The hidden part is more complicated, and cannot be observed through available data, even though it may have a significant influence on the system. The hidden or unobserved part plays a crucial role in governing the overall dynamics of the system [4]. The framework of HMMM allows researchers to study these hidden states indirectly from observable data on employee movements [5].

In past studies, researchers have used HMMM to study workforce systems. However, they have had no conclusive way of deciding how many hidden states to include in the model. Guerry [2] introduced a multinomial Markov-switching manpower model, where the number of hidden states is assumed to be two, arising from two groups of personnel: “movers” and “stayers”. However, [3] increased the number of states to three, including additional group called “mediocre” within the model. Ossai and Madukaife [6] further increased the number of states to five, representing: “high movers”, “movers”, “above mediocre”, “mediocre” and “stayers”. The reseachers were, however, not guided by any specific rule in making the choices of the number of hidden classes in these cases, nor did they check whether these were the best choices.

Commented [HP3]: Please place the referencing appropriately. It should not be just after “However,”. Should it be as “However, Udom and Ebedoro [3]”?

Commented [HP4]: Please remove the extra space.

Commented [HP5]: Please check the spelling, as it should be as “researchers”.

Commented [HP6]: Please simplify the sentences. If possible please make two simple sentences instead of one complex sentence.

Such subjective choice of features of a model can pose significant problems to modelling and applications. When researchers choose inappropriate numbers of hidden states for HMMMs they can arrive at models which have good fits but which may not be the best. It also means that different individuals can specify different HMMMs for the same system. It might make a model less accurate. To this end, [6] proposed a search procedure that uses the likelihood ratio statistic to conduct pairwise comparisons of a series of HMMMs to determine the optimum number of hidden states. This was done for single-department manpower systems, where all the departments are collapsed into one homogeneous structure, using only one metric, the likelihood ratio, for judging the performance of the proposed procedure.

Commented [HP7]: Please place the referencing appropriately.

There are a number of research works outside the area of manpower modelling focused on the problem of choosing the right number of hidden states for the hidden Markov models of interest, see [7, 8, 9]. A common inference from all such works is that the problem is an important one, which has not been satisfactorily resolved, but is tackled according to the requirements of the area of application.

Commented [HP8]: Please add statement(s) to based on the [7, 8, 9] references.

The concept of a general departmentalized manpower framework ($k > 1$ departments) is crucial for dealing with hidden differences within manpower systems. While we can easily break down visible differences like gender into clear categories, there are other hidden differences that might be missed, [1]. Not recognizing them can lead to inaccurate predictions about the system. Despite its importance, very few studies have looked into these hidden differences in manpower systems, likely because it is hard to identify and measure them [2]. So far, only [10] have looked into this hidden differences under a general departmentalized framework. They gave useful insight in the application of HMMM in such framework, underscoring the presence of hidden states in the expanded manpower structure arising from differing propensities for intra-departmental and inter-departmental transitions.

Commented [HP9]: Please remove extra space.

Commented [HP10]: Please add statement(s) to based on the [10] reference.

Commented [HP11]: Please use as “investigated”.

Commented [HP12]: It should be “these”.

In recent years, researchers have shifted focus to modelling workforce systems in general departmentalized frameworks, [11, 12]. The importance of this lies on a truer representation of manpower system dynamics. Ignoring this in modelling the system can cause significant misrepresentation of workforce interactions, [10]. There is also a great need to improve subgroup homogeneity in workforce modelling [13]. Building workforce models in departmentalized framework can be one of the ways of doing this, which also includes how people move between departments [14, 15]. The major penalty in departmentalized workforce modelling is model complexity, which poses no intractable problem in the case of HMMM since, in most cases, it can be handled through the established methods of modelling and estimation of parameters of hidden Markov models.

The focus of the current work is, therefore, on the formulation of HMMMs and procedures for choosing the best number of hidden states out of several alternatives in a general departmentalized manpower framework. The formulations and procedures shall be used to review existing applications of HMMMs on different manpower datasets by researchers in this area.

2. HMMM specification

Consider a manpower system with k observable departments $D_1,...,D_k$, which are internal to the system, and an external department D_{k+1} for individuals who leave to the outside environment. Each $D_i, i=1,...,k+1$, contains u observable categories of members: $g_1,...,g_u$. Thus, d_{is} represents the observable state of being in grade g_s within department D_i. The principle of HMMM assumes that there are hidden states of manpower that governs the workforce flow in each d_{is} according to a Markov chain.

The observable state membership is therefore partitioned by the set of $k \times u$ active states in $D_i, i=1,...,k$ and u non-active states in D_{k+1}. Active state membership ensures probable movement to other states within and outside the active class; non-active state membership means no potential movement to other states within and outside the non-active class. A member of the system must be in one and only one of the states at a time t.

Let the process of state membership in the system be defined by $\{X_t, t \geq 0\}$. Then, $p_{ij(sr)} = P\left(X_{t+1} = d_{jr} \middle| X_t = d_{is}\right)$ is the probability of transition from state d_{is} at time t to d_{jr} at time $t+1$ (from g_s in D_i to g_r in D_j).The following transitions with the corresponding probability and assumptions are possible in the system:

i) Within the active class - intra-departmental transitions: $p_{ii(sr)} = P\left(X_{t+1} = d_{ir} \middle| X_t = d_{is}\right)$,
(assumed possible only for $i=1,...,k; s,r=1,...,u$).

ii) Within the active class - interdepartmental transitions: $p_{ij(sr)} = P\left(X_{t+1} = d_{jr} \middle| X_t = d_{is}\right)$,
(assumed possible only for $i,j=1,...,k; i \neq j; s,r=1,...,u$).

iii) From the active class to non-active class - (transition from g_s in D_i to g_s in D_{k+1}):
$p_{i,k+1(ss)} = P\left(X_{t+1} = d_{k+1,s} \middle| X_t = d_{is}\right)$, (assumed possible only for $i=1,...,k; s,r=1,...,u$).

In general, therefore, the transition probability of all possible transitions in the system can be expressed as: $p_{ij(sr)} = P\left(X_{t+1} = d_{jr} \middle| X_t = d_{is}\right); i,j=1,...,k; s,r=1,...,u,$ where, for intra-departmental transition $j=i$; for inter-departmental transition $i \neq j$; for wastage transition $j=k+1$ and $r=s$. The assumptions above imply: $p_{i,k+1(sr)} = 0$ for all $r \neq s$; $p_{k+1,k+1(sr)} = 0$ for all $r \neq s$; and $p_{k+1,k+1(ss)} = 1$ for all s (all states in D_{k+1} are absorbing).

Let d_{is}, $i=1,...,k$; $s=1,...,u$, have equal number of $k' \geq 1$ hidden states, characterizing workforce flow heterogeneity. Thus, members of each d_{is} are categorized into k', whose transitions to other states are governed by a Markov chain, $\left\{Y_t^{is}, t \geq 0\right\}$. Let $M_l^{is}, l=1,...,k'$, denote the states of the unobserved process $\left\{Y_t^{is}, t \geq 0\right\}$. The $M_l^{is}, l=1,...,k'$, are ranked based on the values of $p_{ij(sr)}$. In all cases, M_1^{is} has the largest transition probability, followed by M_2^{is}, and so on down to $M_{k'}^{is}$. The transition probability of the hidden Markov chain within an observable state d_{is} is denoted by $q_{lm}^{is} = P\left(Y_{t+1}^{is} = M_m^{is} \middle| Y_t^{is} = M_l^{is}\right)$, for $l,m=1,...,k'$, representing the probability of moving from the hidden state M_l^{is} to the hidden state M_m^{is} within the observable state d_{is}. The transition probability of moving from state $M_l^{is}, l=1,...,k'$, in d_{is} to d_{jr} in D_j is given by

Commented [HP13]: Please remove extra space.

Commented [hp14]: Please provide a space.

Commented [hp15]: Please remove extra space.

$p_{ljr}^{is} = P\left(X_{t+1} = d_{jr} \middle| Y_t^{is} = M_l^{is}\right) = P\left(X_{t+1} = d_{jr} \middle| X_t = d_{is}, Y_t^{is} = M_l^{is}\right); l, m = 1, \dots, k'; j = 1, \dots, k;$
$r = 1, \dots, u.$

.

Let the observed number of manpower that transits from d_{is} to d_{jr} within the time interval t to $t+1$ be $n_{ij(sr)}(t)$, so that $n_{ij}(t) = \left(n_{ij(s1)}(t), \dots, n_{ij(su)}(t)\right)$, $j = 1, \dots, k$. Since there are k departments to which manpower can flow, from d_{is}, and since each department has u observable grades, $n_{i(s)}(t) = \left(n_{i1}(t), \dots, n_{ik}(t) n_{i,k+1}(t)\right)$ is the overall manpower transition block vector for state d_{is}, where $n_{i,k+1}(t) = n_{i,k+1(ss)}(t)$ is the number of such flows to the outside environment D_{k+1}.

To establish the probability distribution of the overall flows from d_{is}, $W_{i(s)}(t)$ is defined as the vector having its $j(r)$th member, (for $j = 1, \dots, k+1$; $r = 1, \dots, u$), as $n_{ij(sr)}(t)$. Since $n_{ij(sr)}(t)$ is random, $W_{i(s)}(t)$is random and has a multinomial distribution. This arises from the fact that the event of each transition from d_{is} excludes all other such transitions, and all possible transitions from d_{is} are included in $W_{i(s)}(t)$. Define $Q_{l,n_{i(s)}(t)}^{is}$ as the conditional probability that the multinomial process $W_{i(s)}(t)$ has its value as $n_{i(s)}(t)$, from the interval of time t to $t+1$, given that the state of the chain $\left\{Y_t^{is}, t \geq 0\right\}$ at t is M_l^{is}. That is:

$Q_{l,n_{i(s)}(t)}^{is} = P(W_{i(s)}(t) = n_{i1(s1)}(t), \dots, n_{i1(su)}(t), \dots, n_{ik(s1)}(t), \dots, n_{ik(su)}(t), n_{i,k+1(ss)}(t) \,|\, Y_t^{is} = M_l^{is})$. Thus, $W_{i(s)}(t)$ has the following distribution.

$$Q_{l,n_{i(s)}(t)}^{is} = \binom{N_{i(s)}(t)}{n_{i1(s1)}(t), \dots, n_{i1(su)}(t), \dots, n_{ik(su)}(t), n_{i,k+1(ss)}(t)}$$

$$\times \left(p_{l11}^{is}\right)^{n_{i1(s1)}(t)} \cdots \left(p_{l1u}^{is}\right)^{n_{i1(su)}(t)} \cdots \left(p_{lk1}^{is}\right)^{n_{ik(s1)}(t)} \cdots \left(p_{lku}^{is}\right)^{n_{ik(su)}(t)} \left(p_{l,k+1,s}^{is}\right)^{n_{i,k+1(ss)}(t)}$$

$$\text{for } l = 1, \dots, k'; t = 1, \dots, T \tag{2.1}$$

Where,

$$N_{i(s)}(t) = \sum_{j=1}^{k+1} \sum_{r=1}^{u} n_{ij(sr)}(t) \tag{2.2}$$

For $l = 1, \dots, k'; t = 1, \dots, T$ there are $k'(uk+1)$ probability parameters, p_{ljr}^{is}, in (2.1) to be estimated. The estimates of these parameters are necessary for the implementation of the procedures utilized in this work, for finding the right number of hidden states for a any given manpower system.

3. **Estimation of parameters of HMMM**

The parameters of the HMMM can be estimated using the Maximum Likelihood (ML) estimation method, for simple cases not involving hidden states (or, for cases involving one hidden state), and the Expectation-Maximization (EM) method, for HMMMs with more than one hidden states; see [2, 3, 10]. Rabiner [4] and [16] also provide detailed explanations of EM algorithm in other application areas different from manpower planning. Since interest in the current work is more on the practical implementation of HMMM using the right number of hidden states, only a brief presentation of the estimation methods is done. The probabilities

Commented [hp17]: Please remove the ".".

Commented [hp18]: Should the "th" be the superscript?

Commented [hp19]: Please remove the extra space.

Commented [hp20]: Please provide a space before "is".

Commented [hp21]: Please remove the extra spaces.

Commented [hp22]: Please remove the extra spaces.

Commented [hp23]: Please remove the bold letter from "T".

presented in Section 2 above can be estimated as follows. The ML estimator of $p_{ij(sr)}$, based only on observable state transitions [17], is given, following the preceding definitions, by:

$$\hat{p}_{ij(sr)} = \frac{\sum_{t=1}^{T} n_{ij(sr)}(t)}{\sum_{t=1}^{T} N_{i(s)}(t)} \tag{3.1}$$

The EM estimation for p_{ljr}^{is} and other parameters, which involve hidden states, is more intricate, but tractable. The Expectation step of the EM estimation involves the application of conditional probabilities of the hidden chain, $\{Y_t^{is}\}$, assuming the different hidden states, given the different realizations of the observed values of $W_{i(s)}(t)$ at different time, as well as the forward and backward probabilities of a series of realizations for $W_{i(s)}(t)$ from different standpoint in time. This step leads to the joint likelihood of the observed workforce transitions, $W_{i(s)}(t)$, for $t = 1,\dots,T$; that is:

$$L_T^{is} = P(W_{i(s)}(1) = n_{i1(s1)}(1),\dots,n_{i1(su)}(1),\dots,n_{ik(s1)}(1),\dots,n_{ik(su)}(1),\dots,W_{i(s)}(T) = n_{i1(s1)}(T),$$
$$\dots,n_{i1(su)}(T),\dots,n_{ik(s1)}(T),\dots,n_{ik(su)}(T))$$

By taking the log of L_T^{is} and then the expectation of the resulting log function, which is the end result of the Expectation step, $E\log L_T^{is}$ is obtained as follows in (3.2):

$$\mathrm{ELog}L_T^{is} = \sum_{l=1}^{k'} V_l^{is}(1)\log V_l^{is} + \sum_{t=2}^{T}\sum_{l=1}^{k'}\sum_{m=1}^{k'} U_{l,m}^{is}(t)\log q_{lm}^{is} + \sum_{t=1}^{T}\sum_{l=1}^{k'} V_l^{is}(t)\log Q_{l,n_{i(s)}(t)}^{is} \tag{3.2}$$

Where, $V_l^{is} = P(Y_t^{is} = M_l^{is} \mid W_{i(s)}(1) = n_{i(s)}(1),\dots,W_{i(s)}(T) = n_{i(s)}(T))$, for $l,m = 1,\dots,k'$; V_l^{is} is the initial distribution value of $V_l^{is}(t)$ for the hidden state l; $U_{l,m}^{is}(t) = P(Y_{t-1}^{is} = M_l^{is}, Y_t^{is} = M_m^{is} \mid W_{i(s)}(1) = n_{i(s)}(1),\dots,W_{i(s)}(T) = n_{i(s)}(T))$, for $l,m = 1,\dots,k'$.

In the current application, the maximization step of the EM method involves maximizing $E\log L_T^{is}$ with respect to the constraints on its variables, V_l^{is}, q_{lm}^{is} and $Q_{l,n_{i(s)}(t)}^{is}$. The maximization process, through the Lagrange multipliers method, results to formulas for V_l^{is}, q_{lm}^{is} and $Q_{l,n_{i(s)}(t)}^{is}$ as follows [10].

$$V_l^{is} = V_l^{is}(1), \quad \text{for } l = 1,\dots,k' \tag{3.3}$$

$$q_{lm}^{is} = \frac{\sum_{t=2}^{T} U_{lm}^{is}(t)}{\sum_{m=1}^{k'}\sum_{t=2}^{T} U_{lm}^{is}(t)}, \text{for } l,m = 1,\dots,k' \tag{3.4}$$

$$p_{ljr}^{is} = \frac{\sum_{t=1}^{T} V_l^{is}(t)\,.\,n_{ij(sr)}(t)}{\sum_{t=1}^{T} V_l^{is}(t).N_{i(s)}(t)}, \text{for } l,m = 1,\dots,k';\ j = 1,\dots,k+1;\ r = 1,\dots,u;$$
$$r = s \text{ for } j = k+1 \tag{3.5}$$

The formulas in (3.3), (3.4) and (3.5) are utilized in an iterative manner, for any observed manpower data on which the variables are defined, beginning with a set of initial or chosen

Commented [hp26]: Please remove the extra space.
Commented [hp27]: Please remove the extra space.
Commented [hp28]: Please remove the extra space.
Commented [hp29]: Please remove the extra space.
Commented [hp30]: Please remove the extra space.

starting values, until the values of the parameters converge; that is when (3.2) is maximized. The optimum values obtained become the estimates of the parameters for the HMMM with the given $k^{\acute{}}$. The methods in this section shall be utilized in the procedures for choosing the right number of hidden states of HMMMs in the next section.

4. Procedures for determining the optimum number of hidden states

Two procedures are formulated to admit any statistical test that is appropriate for the problem at hand. In the current study, the Likelihood Ratio statistic (L_r) [18-20], the Akaike Information Criterion (AIC) [21] and the Bayesian Information Criterion (BIC) [22] will be utilized through the procedures to obtain the best $k^{\acute{}}$ for a set of competing HMMMs, on the bases of a desired property of interest, such as model parsimony. The utilization of multiple tests in the search procedures checks result consistency in more than one approach.

Each of the statistic to be employed requires the log likelihood function of the given HMMM. Now, let the HMMM with k' hidden states be HMMM$k^{\acute{}}$ and log likelihood for HMMM$k^{\acute{}}$ be $\log L_{HMMMk^{\acute{}}}$. It can be shown, for the system structure considered in the current work, that for given time t,

$$\log L_{HMMM1} = \sum_{i=1}^{k}\sum_{j=1}^{\mathrm{k}}\sum_{s=1}^{\mathrm{u}}\sum_{r=1}^{\mathrm{u}} n_{ij(sr)}(t)\log p_{ij(sr)} \qquad (4.1)$$

And, for $k^{\acute{}} > 1$,

$$\log L_{HMMMk^{\acute{}}} = \log L_T^{1s} + \ldots + \mathrm{Log} L_T^{ks}$$

$$= \sum_{i=1}^{\mathrm{k}}\Big(\sum_{l=1}^{k^{\acute{}}} V_l^{is}(1)\log \mathrm{P}(Y_1^{is} = M_l^{is} + \sum_{t=2}^{T}\sum_{l=1}^{k^{\acute{}}}\sum_{m=1}^{k^{\acute{}}} U_{l,m}^{is}(t)\log \mathrm{P}(Y_t^{is} = M_m^{is} \big| Y_{t-1}^{is} = M_l^{is})$$
$$+ \sum_{t=1}^{T}\sum_{l=1}^{k^{\acute{}}} V_l^{is}(t)\log \mathrm{P}(W_{i(s)}(t) = n_{i(s)}(t) \big| Y_t^{is} = M_l^{is})) \qquad (4.2)$$

First, a modification of the procedure by [6] to accommodate departmentalization and the implementation through any suitable statistics is done as follows:

Ordered comparison procedure under a desired property:

Step 1: Set $k' = 1,\ldots,N$; the choice of Nmay be guided by prior information about the level of heterogeneity in the system, if any, or any anticipated value can be used depending on available resources.

Step 2: Execute the EM algorithm for all included HMMMk's(for $k' = 1,\ldots,N$) to obtain the estimates of their workforce flow probabilities and check the existence of the models. Eliminate any HMMM that does not exist for the given manpower data.

Step 3: For tests that compare only two models at a time, and where lack of equality of performance is in favour of the HMMM with greater value of k', arrange HMMMk's in decreasing order of a chosen property, such as parsimony, on the basis of k' value; then, conduct pairwise comparison tests following the order of arrangement.

Step 4: For tests that compare only two models at a time, and where lack of equality of performance is in favour of the HMMM with smaller value of k', arrange HMMMk's in

Commented [hp31]: Please remove the extra space.

Commented [hp32]: Please provide a space before "may".

Commented [hp33]: Please rectify the appropriate usage of word.

decreasing order of a chosen property, such as parsimony, on the basis of k' value; then, conduct pairwise comparison tests following the order of arrangement.

Step 5: If Step 3 is followed, identify the pair, following the order, which first records equal performance for the two models; choose the smaller k' of the two models to be the optimum value of k' for the manpower data.

Step 6: If Step 4 is followed, identify the pair, following the order, which first records unequal performance for the two models; choose the smaller k' of the two models to be the optimum value of k' for the manpower data.

Step 7: For tests that can compare more than two models at a time, conduct paired comparison search but reconstruct the current pair according to the decision on the previous pair. Alternatively: compare all the models together as far as the test permits.

Note that the procedure yields an optimum value of k' at the first record of models' equality (following Step 3), or at the first record of models' non equality (following Step 4), since the HMMM with the chosen k' will be preferred to any other succeeding HMMM under the ordered comparison according to the desired property.

Commented [hp34]: Please remove the extra space.

Elimination procedure for paired comparisons:

Step 1: Set $k' = 1, \ldots, N$; the choice of N may be guided by prior information about the level of heterogeneity in the system, if any, or any anticipated value can be used depending on available resources.

Step 2: Execute the EM algorithm for all included HMMM k's (for $k' = 1, \ldots, N$) to obtain the estimates of their workforce flow probabilities and check the existence of the models. Eliminate any HMMM that does not exist for the given manpower data.

Step 3: Test the performance of all included HMMM k's with $k' > 1$ against HMMM1. If non outperforms HMMM1, stop and conclude that there is no significant hidden heterogeneity in the system; that is no point employing HMMM to the system; that is the optimum value of k' is 1. Otherwise, continue with Step 4.

Step 4: If, in Step 3, any of the HMMM k's with $k' > 1$ outperform HMMM1, eliminate HMMM1 and all HMMM k's with $k' > 1$ that failed to outperform HMMM1 from the search. Reconstruct the test for the set of all the HMMMk's with $k' > 1$ that outperformed HMMM1; test all HMMM $k's$ in the set against the HMMM k' in the set with the smallest value of k' and make elimination as before. Continue this process until the last pair. Choose the optimum k' to be the k' of the HMMM that outperforms the other in the last pair.

Note that the set of all the HMMM k's with $k' > 1$ that outperformed HMMM1, as stated in Step 4 above, under elimination procedure for paired comparisons, is where anyone who would not search for optimum value of k' would make an arbitrary choice from. Though such choice would have clear evidence of the need and superiority of the chosen HMMM over HMMM1, it might fall short of the best choice, which is the aim of the procedures.

4.1 Tests for implementing the procedures for identifying the optimum value of k'

The search procedures above admit any suitable tests. For applications in the current work, tests based on L_r, AIC and BIC are employed. For HMMMl and HMMMm, L_r is given as in (4.3):

$$L_r = -2(\log L_{\mathrm{HMMM}_l} - \log L_{\mathrm{HMMM}_m}) \sim \chi^2(\nu) \qquad (4.3)$$

In (4.3), L_r follows chi-square distribution having ν degrees of freedom. Equality of fit of the two models to a given data is rejected in favour of HMMMm if $L_r > \chi^2_\alpha(\nu)$; α is the level of significance.

The AIC compares the goodness of fit of $(N \geq 2)$competing models. It takes model complexity into consideration through the number of parameters, p. The model with the smallest value of AIC is judged the best for a given data. The AIC measure for HMMMl is given in (4.4), see [23].

$$\mathrm{AIC} = 2p - 2\log L_{\mathrm{HMMM}l} \tag{4.4}$$

Similarly, BIC compares $(N \geq 2)$ competing models in favour of the model with the smallest value, with a higher penalty against models with more parameters. The BIC measure for HMMMl, with number of observations w, is given in (4.5).

$$\mathrm{BIC} = p\log w - 2\log L_{\mathrm{HMMM}l} \tag{4.5}$$

The use of L_r requires computing log likelihoods and ν. The log likelihoods are computed from stages of the EM algorithm. The value of ν in (4.3) is computed as the difference between the free parameters of the two models. For the general case of departmentalized system in this work, the following result for ν, Result 4.1, covers all cases of pairwise comparison under L_r test.

Result 4.1
For a general departmentalized system with k departments and u categories per department, ν for L_r test in (4.3), for $m \geq l$, is given as follows:

$$\nu = uk(m - l\ )(uk + m + l)$$

Proof
Given HMMMl, the free parameters are obtained thus: there are l hidden classes in each state, d_{is}, from where workforce flows to other $uk+1$states in and outside the system, resulting to luk free parameters. Also, the system dynamics within d_{is}, through the l hidden states, results to $(l-1)$ $l$free transition probability parameters and $(l-1)$ initial state probability parameters. These give, altogether, $luk + (l-1)l + (l-1)$ free parameters from each d_{is} in HMMMl. So, overall, there are $uk(luk + (l-1)l + (l-1))$ free parameters in HMMMl. Similarly, there are $uk(muk + (m-1)m + (m-1))$ free parameters in HMMMm. The result follows from the definition of ν.

In the application of the search procedures in the current paper, N is to be set at 5; so, $k' = 1, \ldots, 5$. Table 1, therefore, contains all the formulas that lead to the realization of all the values of ν that may be of use when $N \leq 5$.

Table 1: Specific Formulas for ν in the L_r Test in Search Procedures with $N \leq 5$

Models Compared	Formula for ν
HMMM1 vs. HMMM2	$(uk)^2 + 3uk$
HMMM1 vs. HMMM3	$2(uk)^2 + 8uk$
HMMM1 vs. HMMM4	$3(uk)^2 + 15uk$
HMMM1 vs. HMMM5	$4(uk)^2 + 24uk$
HMMM2 vs. HMMM3	$(uk)^2 + 5uk$

Commented [HP35]: Please provide a space before "competing".

Commented [HP36]: Please provide a sentence to refer the references.

Commented [HP37]: Please remove the extra space.

Commented [HP38]: Please add a space before "states".

Commented [HP39]: Please add a space before "free".

HMMM2 vs. HMMM4	$2(uk)^2+12uk$
HMMM2 vs. HMMM5	$3(uk)^2+21uk$
HMMM3 vs. HMMM4	$(uk)^2+7uk$
HMMM3 vs. HMMM5	$2(uk)^2+16uk$
HMMM4 vs. HMMM5	$(uk)^2+9uk$

5. Numerical application

The techniques discussed in the previous sections will be applied to manpower data used by different researchers in the area of hidden Markov manpower model. This move has two main objectives. The first is to illustrate how the optimum number of hidden states in hidden Markov manpower models can be identified using the procedures. The second is to review the work of these researchers to ascertain whether they chose the right number of hidden states. Specifically, the L_r, AIC and BIC tests will be utilized in the procedures to re-analyze data used by [2], [3] and [10].

The dataset used by [10] is that of a university non-academic manpower system, which is divided into two main departments: the academic department (AD), corresponding to i or j = 1 , and the non-academic department (NAD), corresponding to i or j = 2. The personnel within these departments are classified into two main categories: junior staff (JS), corresponding to s or $r=1$, and senior staff (SS), corresponding to s or $r=2$. In the study conducted by [3], a single-department university academic manpower data, previously analyzed by [24], were utilized. The data contain the academic staff record of a university, where the staff are categorized in ascending order of rank: graduate assistant (G_1), assistant lecturer (G_2), lecturer II (G_3), lecturer I (G_4), senior lecturer (G_5), associate professor (G_6) and professor (G_7). An additional category (G_8), is also included to represent staff leaving the system (wastage). Also, [2] analyzed the personnel data from an organization over a nine-year period, collecting information on various employee characteristics such as job grade, monthly working hours, age, years of service, and number of children. Using this dataset, it was possible to note internal staff movements such as promotions and transfers as well as staff wastage, including retirements and resignations.

For the purpose of application in the current work, for each dataset, going by the proposed procedures, the existence of the HMMM k' sfor all k' (where $k'=1,\ldots,5$) is first assumed. This assumption is confirmed by estimating the parameters of each model using the EM formulas. The resulting transition probabilities are then analyzed to confirm that the hidden classes are effectively represented by distinct transition probabilities. In the case of [10], this confirmation has already been done for all $k'=1,\ldots,5$, where also $k'=2$ was chosen as the optimum value, but on the basis of L_r procedure only. This is reviewed here on the basis of all the search procedures presented in the current work. Likewise, [2] confirmed the existence of HMMM1 and HMMM3 for the manpower system they studied, but only compared these two models and chose HMMM3 as the better one. This is reviewed here on the basis of the search procedures presented in the current work, for $k'=1,\ldots,5$. Also, [3] confirmed and compared HMMM1 and HMMM2 for the system studied, and chose HMMM2 as the better model on the basis of L_r test only.

Therefore, to confirm the existence of HMMM k' for all $k'=1,\ldots,5$, the estimate of the transition probabilities is computed only for $k'=2,4,5$ for the data of [2] and for $k'=3,4,5$ for the data of [3]. The transition probability matrix (tpm) for each of the five models HMMM1, HMMM2, HMMM3, HMMM4and HMMM5 are respectively denoted as P_1, P_2, P_3, P_4 and P_5

Commented [HP40]: Please remove the extra space.

Commented [HP41]: Please remove the extra space.

Commented [HP42]: Please provide a space before "for".

Commented [HP43]: Please provide a space before "and".

for data of [10]; as $p_1{}^U$, $p_2{}^U$, $p_3{}^U$,$p_4{}^U$ and $p_5{}^U$ for data of [2]; and as $p_1{}^G$, $p_2{}^G$, $p_3{}^G$,$p_4{}^G$and $p_5{}^G$ for data of [3]. These matrices are shown below, (the omitted ones are contained in the respective works cited). Tables 2-7 contain the results of the search procedures for the data utilized.

Commented [HP44]: Did not see the TPM for $P_1{}^U$, $P_1{}^G$, $P_2{}^G$?

$$P_2^U = \begin{array}{cc} & \begin{array}{cccccccc} G_1 & G_2 & G_3 & G_4 & G_5 & G_6 & G_7 & G_8 \end{array} \\ \begin{array}{c} G_1 \\ \\ G_2 \\ \\ G_3 \\ \\ G_4 \\ \\ G_5 \\ \\ G_6 \\ \\ G_7 \\ \\ \end{array} & \begin{pmatrix} 0.963 & 0.021 & 0 & 0 & 0 & 0 & 0 & 0.016 \\ 0.991 & 0 & 0 & 0 & 0 & 0 & 0 & 0.009 \\ 0 & 0.931 & 0.060 & 0 & 0 & 0 & 0 & 0.009 \\ 0 & 0.969 & 0.026 & 0 & 0 & 0 & 0 & 0.005 \\ 0 & 0 & 0.935 & 0.047 & 0 & 0 & 0 & 0,018 \\ 0 & 0 & 0.965 & 0.021 & 0 & 0 & 0 & 0.014 \\ 0 & 0 & 0 & 0.941 & 0.057 & 0 & 0 & 0.002 \\ 0 & 0 & 0 & 0.921 & 0.069 & 0 & 0 & 0.010 \\ 0 & 0 & 0 & 0 & 0.960 & 0.030 & 0 & 0.011 \\ 0 & 0 & 0 & 0 & 0.968 & 0.030 & 0 & 0.002 \\ 0 & 0 & 0 & 0 & 0 & 0.962 & 0.023 & 0.015 \\ 0 & 0 & 0 & 0 & 0 & 0.918 & 0.045 & 0.036 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.984 & 0.016 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.967 & 0.033 \end{pmatrix} \end{array}$$

$$P_4^U = \begin{array}{cc} & \begin{array}{cccccccc} G_1 & G_2 & G_3 & G_4 & G_5 & G_6 & G_7 & G_8 \end{array} \\ \begin{array}{c} \\ G_1 \\ \\ \\ \\ G_2 \\ \\ \\ \\ G_3 \\ \\ \\ \\ G_4 \\ \\ \\ \\ G_5 \\ \\ \\ \\ G_6 \\ \\ \\ \\ G_7 \\ \\ \\ \end{array} & \begin{pmatrix} 0.963 & 0.021 & 0 & 0 & 0 & 0 & 0 & 0.016 \\ 0.983 & 0 & 0 & 0 & 0 & 0 & 0 & 0.017 \\ 1.000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.963 & 0.021 & 0 & 0 & 0 & 0 & 0 & 0.016 \\ 0 & 0.925 & 0.067 & 0 & 0 & 0 & 0 & 0.008 \\ 0 & 0.935 & 0.054 & 0 & 0 & 0 & 0 & 0.011 \\ 0 & 0.969 & 0.026 & 0 & 0 & 0 & 0 & 0.005 \\ 0 & 0.925 & 0.067 & 0 & 0 & 0 & 0 & 0.008 \\ 0 & 0 & 0.949 & 0.035 & 0 & 0 & 0 & 0.016 \\ 0 & 0 & 0.932 & 0.051 & 0 & 0 & 0 & 0.017 \\ 0 & 0 & 0.965 & 0.021 & 0 & 0 & 0 & 0.014 \\ 0 & 0 & 0.923 & 0.055 & 0 & 0 & 0 & 0.022 \\ 0 & 0 & 0 & 0.927 & 0.073 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.921 & 0.069 & 0 & 0 & 0.010 \\ 0 & 0 & 0 & 0.953 & 0.044 & 0 & 0 & 0.003 \\ 0 & 0 & 0 & 0.927 & 0.073 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.960 & 0.030 & 0 & 0.011 \\ 0 & 0 & 0 & 0 & 0.973 & 0.027 & 0 & 0.000 \\ 0 & 0 & 0 & 0 & 0.962 & 0.034 & 0 & 0.003 \\ 0 & 0 & 0 & 0 & 0.960 & 0.030 & 0 & 0.011 \\ 0 & 0 & 0 & 0 & 0 & 0.962 & 0.026 & 0.012 \\ 0 & 0 & 0 & 0 & 0 & 0.962 & 0.019 & 0.019 \\ 0 & 0 & 0 & 0 & 0 & 0.918 & 0.045 & 0.036 \\ 0 & 0 & 0 & 0 & 0 & 0.962 & 0.026 & 0.012 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.986 & 0.014 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.980 & 0.020 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.967 & 0.033 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.986 & 0.014 \end{pmatrix} \end{array}$$

$$
P_5^U = \begin{array}{c} \\ G_1 \\ \\ \\ \\ \\ \\ G_2 \\ \\ \\ \\ \\ G_3 \\ \\ \\ \\ \\ G_4 \\ \\ \\ \\ \\ G_5 \\ \\ \\ \\ \\ G_6 \\ \\ \\ \\ \\ G_7 \\ \\ \\ \end{array}
\begin{array}{c}
\begin{array}{cccccccc} G_1 & G_2 & G_3 & G_4 & G_5 & G_6 & G_7 & G_8 \end{array} \\
\left(\begin{array}{cccccccc}
0.959 & 0.031 & 0 & 0 & 0 & 0 & 0 & 0.010 \\
0.967 & 0.011 & 0 & 0 & 0 & 0 & 0 & 0.022 \\
0.983 & 0 & 0 & 0 & 0 & 0 & 0 & 0.017 \\
1.000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.959 & 0.031 & 0 & 0 & 0 & 0 & 0 & 0.010 \\
0 & 0.915 & 0.079 & 0 & 0 & 0 & 0 & 0.006 \\
0 & 0.935 & 0.055 & 0 & 0 & 0 & 0 & 0.011 \\
0 & 0.941 & 0.048 & 0 & 0 & 0 & 0 & 0.011 \\
0 & 0.969 & 0.026 & 0 & 0 & 0 & 0 & 0.005 \\
0 & 0 & 0.079 & 0 & 0 & 0 & 0 & 0.006 \\
0 & 0 & 0.949 & 0.035 & 0 & 0 & 0 & 0.016 \\
0 & 0 & 0.923 & 0.055 & 0 & 0 & 0 & 0.022 \\
0 & 0 & 0.932 & 0.051 & 0 & 0 & 0 & 0.017 \\
0 & 0 & 0.965 & 0.021 & 0 & 0 & 0 & 0.014 \\
0 & 0 & 0.949 & 0.035 & 0 & 0 & 0 & 0.016 \\
0 & 0 & 0 & 0.927 & 0.073 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.954 & 0.043 & 0 & 0 & 0.003 \\
0 & 0 & 0 & 0.913 & 0.078 & 0 & 0 & 0.009 \\
0 & 0 & 0 & 0.931 & 0.059 & 0 & 0 & 0.010 \\
0 & 0 & 0 & 0.927 & 0.073 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.957 & 0.033 & 0 & 0.011 \\
0 & 0 & 0 & 0 & 0.964 & 0.026 & 0 & 0.011 \\
0 & 0 & 0 & 0 & 0.962 & 0.034 & 0 & 0.003 \\
0 & 0 & 0 & 0 & 0.973 & 0.027 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.957 & 0.033 & 0 & 0.011 \\
0 & 0 & 0 & 0 & 0 & 0.957 & 0.034 & 0.009 \\
0 & 0 & 0 & 0 & 0 & 0.969 & 0.016 & 0.016 \\
0 & 0 & 0 & 0 & 0 & 0.960 & 0.020 & 0.020 \\
0 & 0 & 0 & 0 & 0 & 0.918 & 0.045 & 0.036 \\
0 & 0 & 0 & 0 & 0 & 0.957 & 0.034 & 0.009 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.986 & 0.014 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.986 & 0.014 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.980 & 0.020 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.967 & 0.030 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.986 & 0.014
\end{array}\right)
\end{array}
$$

$$
P_3^G = \begin{array}{c} \\ H_1 \\ \\ \\ \\ H_2 \\ \\ \\ H_3 \\ \\ \\ H_4 \\ \\ \end{array}
\begin{array}{c}
\begin{array}{ccccc} H_1 & H_2 & H_3 & H_4 & H_9 \end{array} \\
\left(\begin{array}{ccccc}
0.493 & 0.031 & 0.415 & 0.031 & 0.031 \\
0.908 & 0.018 & 0.038 & 0.018 & 0.018 \\
0.784 & 0.028 & 0.132 & 0.028 & 0.028 \\
0.040 & 0.840 & 0.040 & 0.040 & 0.040 \\
0.014 & 0.929 & 0.014 & 0.030 & 0.014 \\
0.025 & 0.897 & 0.025 & 0.028 & 0.025 \\
0 & 0 & 0.410 & 0.571 & 0.019 \\
0 & 0 & 0.111 & 0.872 & 0.016 \\
0 & 0 & 0.230 & 0.751 & 0.019 \\
0 & 0 & 0.350 & 0.631 & 0.020 \\
0 & 0 & 0.364 & 0.615 & 0.021 \\
0 & 0 & 0.444 & 0.533 & 0.022
\end{array}\right)
\end{array}
$$

$$
P_4^G = \begin{array}{cc}
 & \begin{array}{ccccc} H_1 & H_2 & H_3 & H_4 & H_9 \end{array} \\
\begin{array}{c} \\ H_1 \\ \\ \\ \\ H_2 \\ \\ \\ \\ H_3 \\ \\ \\ \\ H_4 \\ \\ \\ \end{array} &
\begin{pmatrix}
0.492 & 0.031 & 0.415 & 0.031 & 0.031 \\
0.905 & 0.018 & 0.040 & 0.018 & 0.018 \\
0.767 & 0.033 & 0.133 & 0.033 & 0.033 \\
0.763 & 0.026 & 0.158 & 0.026 & 0.026 \\
0.040 & 0.840 & 0.040 & 0.040 & 0.040 \\
0.029 & 0.886 & 0.029 & 0.029 & 0.029 \\
0.013 & 0.926 & 0.013 & 0.031 & 0.013 \\
0.022 & 0.911 & 0.022 & 0.022 & 0.022 \\
0 & 0 & 0.355 & 0.629 & 0.016 \\
0 & 0 & 0.488 & 0.488 & 0.023 \\
0 & 0 & 0.220 & 0.761 & 0.019 \\
0 & 0 & 0.110 & 0.873 & 0.016 \\
0 & 0 & 0.314 & 0.667 & 0.020 \\
0 & 0 & 0.444 & 0.533 & 0.022 \\
0 & 0 & 0.383 & 0.593 & 0.025 \\
0 & 0 & 0.358 & 0.623 & 0.020
\end{pmatrix}
\end{array}
$$

$$
P_5^G = \begin{array}{cc}
 & \begin{array}{ccccc} H_1 & H_2 & H_3 & H_4 & H_9 \end{array} \\
\begin{array}{c} \\ H_1 \\ \\ \\ \\ \\ H_2 \\ \\ \\ \\ \\ \\ H_3 \\ \\ \\ \\ \\ H_4 \\ \\ \\ \end{array} &
\begin{pmatrix}
0.918 & 0.017 & 0.031 & 0.017 & 0.017 \\
0.763 & 0.026 & 0.158 & 0.026 & 0.026 \\
0.492 & 0.031 & 0.415 & 0.031 & 0.031 \\
0.868 & 0.022 & 0.065 & 0.022 & 0.022 \\
0.767 & 0.033 & 0.133 & 0.033 & 0.033 \\
0.013 & 0.932 & 0.013 & 0.030 & 0.013 \\
0.018 & 0.911 & 0.018 & 0.036 & 0.018 \\
0.040 & 0.840 & 0.040 & 0.040 & 0.040 \\
0.022 & 0.911 & 0.022 & 0.022 & 0.022 \\
0.029 & 0.886 & 0.029 & 0.029 & 0.029 \\
0 & 0 & 0.109 & 0.875 & 0.016 \\
0 & 0 & 0.255 & 0.725 & 0.020 \\
0 & 0 & 0.488 & 0.488 & 0.023 \\
0 & 0 & 0.198 & 0.784 & 0.018 \\
0 & 0 & 0.355 & 0.629 & 0.016 \\
0 & 0 & 0.359 & 0.621 & 0.020 \\
0 & 0 & 0.355 & 0.625 & 0.020 \\
0 & 0 & 0.444 & 0.533 & 0.022 \\
0 & 0 & 0.383 & 0.593 & 0.025 \\
0 & 0 & 0.314 & 0.667 & 0.020
\end{pmatrix}
\end{array}
$$

Table 2: Results of ordered comparison search procedure for data of [10]

Compared models	Log-Likelihood Value	L_r	v; $\chi^2_{0.05}(v)$	AIC	BIC	Better model based on L_r	Chosen model based on AIC	Chosen model based on BIC	Best model based on L_r, AIC & BIC
HMMM1 vs. HMMM2	-210.05 (-155.87)	108.36	28, 41.34	468.10 (407.75)	467.68 (406.90)	HMMM2	HMMM2	HMMM2	HMMM2
HMMM1 vs. HMMM3	-210.05 (-138.97)	142.15	64, 79.08	468.10 (421.95)	467.68 (420.67)	HMMM3	HMMM3	HMMM3	HMMM3
HMMM1 vs. HMMM4	-210.05 (-132.81)	154.49	108, 138.63	468.10 (457.61)	467.68 (455.91)	HMMM4	HMMM4	HMMM4	HMMM4
HMMM1 vs. HMMM5	-210.05 (-119.93)	180.24	160, 201.60	468.10 (479.86)	467.68 (477.73)	Equal Fit	HMMM1	HMMM1	HMMM1
HMMM2 vs. HMMM3	-155.87 (-138.97)	33.80	36, 51.00	407.75 (421.95)	406.90 (420.67)	Equal Fit			
HMMM3 vs. HMMM4	-138.97 (-132.80)	12.33	44, 67.50	421.95 (457.61)	420.67 (455.91)	Equal Fit	HMMM2	HMMM2	HMMM2
HMMM4 vs. HMMM5	-132.80 (-119.93)	25.75	52, 79.08	457.61 (479.86)	455.91 (477.73)	Equal Fit			

(Note: The value in bracket corresponds to the second model in each comparison)

Table 3: Results of ordered comparison search procedure for data of [2]

Compared models	Log-Likelihood Value	L_r	v; $\chi^2_{0.05}(v)$	AIC	BIC	Better model based on L_r	Chosen model based on AIC	Chosen model based on BIC	Best model based on L_r, AIC & BIC
HMMM1vs. HMMM2	-533.82 (-123.56)	820.52	70, 87.50	1179.64 (471.12)	1199.25 (510.35)	HMMM2	HMMM2	HMMM2	HMMM2
HMMM1 vs. HMMM3	-533.82 (-84.01)	899.62	154, 193.98	1179.64 (504.02)	1199.25 (562.86)	HMMM3	HMMM3	HMMM3	HMMM3
HMMM1 vs. HMMM4	-533.82 (-80.75)	906.14	252, 293.68	1179.64 (609.50)	1199.25 (687.96)	HMMM4	HMMM4	HMMM4	HMMM4
HMMM1 vs. HMMM5	-533.82 (-76.85)	913.94	364, 407.46	1179.64 (713.70)	1199.25 (811.77)	HMMM5	HMMM5	HMMM5	HMMM5
HMMM2 vs. HMMM3	-123.56 (-84.01)	79.10	84, 106.40	471.12 (504.02)	510.35 (562.86)	Equal Fit			
HMMM3 vs. HMMM4	-84.01 (-80.75)	6.52	98, 122.11	504.02 (609.50)	420.6705 (687.96)	Equal Fit			
HMMM4 vs. HMMM5	-80.75 (-76.85)	7.80	112, 136.14	609.50 (713.70)	687.96 (811.77)	Equal Fit	HMMM2	HMMM2	HMMM2

(Note: The value in bracket corresponds to the second model in each comparison)

Table 4: Results of ordered comparison search procedure for data of [3]

Compared models	Log-Likelihood Value	L_r	v; $\chi^2_{0.05}(v)$	AIC	BIC	Better model based on L_r	Chosen model based on AIC	Chosen model based on BIC	Best model based on L_r, AIC & BIC
HMMM1 vs. HMMM2	-199.07 (-155.61)	86.93	28, 41.34	438.14 (391.21)	443.24 (401.42)	HMMM2	HMMM2	HMMM2	HMMM2
HMMM1 vs. HMMM3	-199.07 (-148.06)	102.01	64, 79.08	438.14 (416.13)	443.24 (431.44)	HMMM3	HMMM3	HMMM3	HMMM3
HMMM1 vs. HMMM4	-199.07 (-144.07)	109.99	108, 138.63	438.14 (448.15)	443.24 (468.57)	Equal Fit	HMMM1	HMMM1	HMMM1
HMMM1 vs. HMMM5	-199.07 (-141.95)	114.24	160, 201.60	438.14 (483.90)	443.24 (509.42)	Equal Fit	HMMM1	HMMM1	HMMM1
HMMM2 vs. HMMM3	-155.61 (-148.06)	15.08	36, 51.00	391.21 (416.13)	401.42 (431.44)	Equal Fit	HMMM2	HMMM2	HMMM2
HMMM3 vs. HMMM4	-148.06 (-144.073)	7.98	44, 67.50	416.13 (448.15)	431.44 (468.57)	Equal Fit			
HMMM4 vs. HMMM5	-144.07 (-141.95)	4.25	52, 79.08	448.15 (483.90)	468.57 (509.42)	Equal Fit			

(Note: The value in bracket corresponds to the second model in each comparison)

Table 5: Results of elimination search procedure for data of [10]

Stage	Compared models	Log-Likelihood Value	L_r	v; $\chi^2_{0.05}(v)$	AIC	BIC	Eliminated models based on L_r	Eliminated models based on AIC	Eliminated models based on BIC	Best model based on L_r, AIC & BIC
1	HMMM1 vs. HMMM2	-210.05 (-155.87)	108.36	28, 41.34	468.10 (407.75)	467.68 (406.90)	HMMM1 HMMM5	HMMM1 HMMM5	HMMM1 HMMM5	Not yet found
	HMMM1 vs. HMMM3	-210.05 (-138.97)	142.15	64, 79.08	468.10 (421.95)	467.68 (420.67)				
	HMMM1 vs. HMMM4	-210.05 (-132.81)	154.49	108, 138.63	468.10 (457.61)	467.68 (455.91)				
	HMMM1 vs. HMMM5	-210.05 (-119.93)	180.24	160, 201.60	468.10 (479.86)	467.68 (477.73)				
2	HMMM2 vs. HMMM3	-155.87 (-138.97)	33.80	36, 51.00	407.75 (421.95)	406.90 (420.67)	HMMM3 HMMM4	HMMM3 HMMM4	HMMM3 HMMM4	HMMM2
	HMMM2 vs. HMMM4	-155.87 (-132.80)	46.14	80, 101.89	407.75 (457.61)	406.90 (455.91)				

(Note: The value in bracket correspond to the second model in each comparison)

Table 6: Results of elimination search procedure for data of [2]

Stage	Compared models	Log-Likelihood Value	L_r	v; $\chi^2_{0.05}(v)$	AIC	BIC	Eliminated models based on L_r	Eliminated models based on AIC	Eliminated models based on BIC	Best model based on L_r, AIC & BIC
1	HMMM1 vs. HMMM2	-533.82 (-123.56)	820.52	70, 87.50	1179.64 (471.12)	1199.25 (510.35)	HMMM1	HMMM1	HMMM1	Not yet found
	HMMM1 vs. HMMM3	-533.82 (-84.01)	899.62	154, 193.98	1179.64 (504.02)	1199.25 (562.86)				
	HMMM1 vs. HMMM4	-533.82 (-80.75)	906.14	252, 293.68	1179.64 (609.50)	1199.25 (687.96)				
	HMMM1 vs. HMMM5	-533.82 (-76.85)	913.94	364, 407.46	1179.64 (713.70)	1199.25 (811.77)				
2	HMMM2 vs. HMMM3	-123.56 (-84.01)	79.10	84, 106.40	471.12 (504.02)	510.35 (562.86)	HMMM3 HMMM4 HMMM5	HMMM3 HMMM4 HMMM5	HMMM3 HMMM4 HMMM5	HMMM2
	HMMM2 vs. HMMM4	-123.56 (-80.75)	85.62	182, 225.56	471.12 (609.50)	510.35 (687.96)				
	HMMM2 vs. HMMM5	-123.56 (-76.85)	93.42	294, 358.58	471.12 (713.70)	510.35 (811.77)				

(Note: The value in bracket correspond to the second model in each comparison)

Table 7: Results of elimination search procedure for data of [3]

Stage	Compared models	Log-Likelihood Value	L_r	v; $\chi^2_{0.05}(v)$	AIC	BIC	Eliminated models based on L_r	Eliminated models based on AIC	Eliminated models based on BIC	Best model based on L_r, AIC & BIC
1	HMMM1 vs. HMMM2	-199.07 (-155.61)	86.93	28, 41.34	438.14 (391.21)	443.24 (401.42)	HMMM1 HMMM4 HMMM5	HMMM1 HMMM4 HMMM5	HMMM1 HMMM4 HMMM5	Not yet found
	HMMM1 vs. HMMM3	-199.07 (-148.06)	102.01	64, 79.08	438.14 (416.13)	443.24 (431.44)				
	HMMM1 vs. HMMM4	-199.07 (-144.07)	109.99	108, 138.63	438.14 (448.15)	443.24 (468.57)				
	HMMM1 vs. HMMM5	-199.07 (-141.95)	114.24	160, 201.60	438.14 (483.90)	443.24 (509.42)				
2	HMMM2 vs. HMMM3	-155.61 (-148.06)	15.08	36, 51.00	391.21 (416.13)	401.42 (431.44)	HMMM3	HMMM3	HMMM3	HMMM2

(Note: The value in bracket correspond to the second model in each comparison)

5.1 Discussion of results

In order to implement the search procedures for the manpower data analyzed in this work, the first two steps are, summarily, to chose the number of HMMMs, *N*, to be included, and then to investigate their existence. These have been done for the three datasets, where *N* was chosen to be 5 in each case. Also, all the included HMMMs: HMMM1, HMMM2, HMMM3, HMMM4 and HMMM5, exist in each case, as seen from their estimated manpower transition probabilities. In other words, all the estimated probabilities in the five transition probability matrices corresponding to the five HMMMs exist and are well defined in the three cases reviewed. For instance, consider the matrices ${p_2}^U$ and ${p_5}^U$. The value, ${p_2}^U(2,3)$ denotes the probability of an assistant lecturer moving to lecturer II. But, ${p_2}^U(2,3)$ has two values, 0.060 and 0.026, in ${p_2}^U$, indicating that two hidden states exist, which influence this transition differently. Members of the first hidden state have a higher probability, 0.060, than the second state, 0.026, of making this same transition. In the case of ${p_5}^U$, the same probability of an assistant lecturer moving to lecturer II, ${p_5}^U(2,3)$, now has five probability values, 0.079, 0.055, 0.048, 0.026 and 0.079, indicating that five hidden states are also possible, which influence this transition differently. The value 0.079 being repeated is still in order, since the existence is holistically considered row wise. Similar results for existence of the models can also be observed for all the other states of the manpower systems under review using the estimated transition probabilities. All indicate that all the models ($k'=1,...,5$) are valid and feasible representations of the manpower data being analyzed.

Further steps of the ordered comparison and elimination search procedures for the three datasets in Tables 2, 3, 4, 5, 6 and 7 show strikingly consistent results. The main result is that in all the cases (datasets reviewed), $k'=2$ is the optimum number of hidden state. That is, HMMM2 is the best model among the five HMMMs in all the cases. In evaluating the ordered comparison search procedure for the three datasets, row 2 and row 6 of Tables 2, 3 and 4 are sufficient for making conclusion. This is because, in each of these three tables, at row 2, HMMM2 outperformed HMMM1 based on all of the three tests, L_r, AIC and BIC, (for instance, in row 2 in Table 4, L_r=86.93 $>\chi^2_{0.05}(28)$=41.34 and both the AIC and BIC values, 438.14 and 443.24, are respectively greater than those for HMMM2,391.21 and 401.42). Then, going higher in the ordered models, HMMM2 is compared with HMMM3 (row 6), where both models perform equally based on L_r test and HMMM2 outperformed HMMM3 based on AIC and BIC. This (in row 6) is the first point of equal performance along the ordered models, and it is where the search stops and the smaller value of k' chosen as optimum. The inclusion of other rows in Tables 2, 3 and 4 are a sort of sensitivity analysis on what could happen around the main steps of the procedure, which, interestingly, further validate the main result.

Also, the results that led to $k'=2$ being the optimum number of hidden state in all the three cases are evident from the elimination search procedure in Tables 5, 6 and 7. In Table 5, HMMM1 and HMMM5 are eliminated in stage 1, while HMMM3 and HMMM4 are eliminated in stage 2, based on all the three tests, L_r, AIC and BIC, leaving HMMM2 as the model with the best value of k'. Similarly, in Table 6 HMMM1 is eliminated in stage 1, while HMMM3, HMMM4 and HMMM5 are eliminated in stage 2, based on all the three tests, L_r, AIC and BIC, leaving HMMM2 as the model with the best value of k'. Again, in Table 7, HMMM1, HMMM4 and HMMM5 are eliminated in stage 1, while HMMM3 is eliminated in stage 2, based on all the three tests, L_r, AIC and BIC, leaving HMMM2 as the model with the best value of k'.

6. Conclusion

In this article, systematic approaches for determining the optimum number of hidden states in a hidden Markov manpower model were established in a more general departmentalized framework. Traditionally, the choice of the number of hidden states in Markov manpower models has been subjective, as observable data cannot readily afford a concrete way of doing this. The two approaches employed to address this, the ordered comparison and elimination search procedures, can consistently entertain different suitable statistical tests, and are, therefore, robust for this purpose.

The application of the introduced approaches on the three manpower datasets used in the area of hidden Markov manpower model shows a striking consistency in pointing out the optimum number of hidden states. In all the three datasets reviewed, $k' = 2$ is shown to be the optimum number of hidden states. This, unfortunately, was not the arbitrary choice of all the three researchers who had previously analyzed the datasets. Only two of them made this right choice.

The results also show that increasing the number of hidden states beyond a threshold, which is $k' = 2$ for the three manpower datasets used in the current work, may not significantly improve the goodness of fit of the model. Rather, it undermines model parsimony by adding complexity without commensurate enhancement in fit.

Furthermore, it is evident from the results from model comparison in this work that almost all the models with $k' \geq 2$ outperformed HMMM1 in all the cases considered. This supports the conclusion that hidden heterogeneity is inherent in manpower systems. It also supports the argument of the current work, that the number of hidden states should not be chosen arbitrarily.

For further work, a single metric or statistical tool that can combine all the procedures in one for judging the best number of hidden states of a HMMM can be sort for. This can save from the tedious work of calculating each component of the procedures separately. Also, since a common point, $k' = 2$, beyond which model fit does not improve, was found for the three datasets from manpower systems of different settings and workforce, there may exist such threshold point or set of points that may be general for manpower systems; such may require simulation. Also, it is worthy of note that the approaches proposed in this work, for identifying the right number of hidden state of a hidden Markov model, can readily be adapted and applied in other areas outside manpower systems, as long as the system conditions are similar.

Commented [HP47]: Please remove the extra space.

Commented [HP48]: Should this be "states"?

Competing Interests

Authors declared that no competing interests exist.

References

Commented [HP49]: If at least couple of latest (may be from last 2-3 years) research papers to be included.

[1] Ugwuowo, F. I. & McClean, S. I. (2000). Modelling heterogeneity in a manpower system: A review. Applied Stochastic Models in Business and Industry, 16(2), 99-110. https://doi.org/10.1002/1526-4025(200004/06)16:2<99::AID-ASMB385>3.0.CO;2-3

[2] Udom, A. U. & Ebedoro, U. G. (2021). On multinomial hidden Markov model for hierarchical manpower systems. Communications in Statistics-Theory and Methods, 50(6), 1370-1386. https://doi.org/10.1080/03610926.2019.1650185

[3] Guerry, M. A. (2011). Hidden heterogeneity in manpower systems: A Markov-switching

model approach. European Journal of Operational Research, 210(1), 106-113. https://doi.org/10.1016/j.ejor.2010.10.039

[4] Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2), 257-286. https://doi.org/10.1109/5.18626

[5] Bartholomew, D. J. (1982). Stochastic models for social processes (3rd ed). Chichester: Wiley.

[6] Ossai, E. O. & Madukaife, M. S. (2022). Likelihood ratio search procedure for optimum number of states in a hidden Markov manpower model", Asian Journal of Probability and Statistics, 20(4), 146-156. https://doi.org/10.9734/ajpas/2022/v20i4445

[7] Pohle, J., Langrock, R., van Beest, F.M. & Schmidt, N. M. (2017). Selecting the number of states in hidden Markov models: pragmatic solutions illustrated using animal movement. JABES, 22, 270–293. https://doi.org/10.1007/s13253-017-0283-8

[8] Lin, Y. & Song, X. (2022). Order selection for regression-based hidden Markov model. Journal of Multivariate Analysis, 192. https://doi.org/10.1016/j.jmva.2022.105061

[9] Buckby, J., Wang, T., Fletcher, D., Zhuang, J., Takeo, A. & Obara, K. (2023). Finding the number of latent states in hidden Markov models using information criteria. Environ Ecol Stat, 30, 797–825. https://doi.org/10.1007/s10651-023-00584-5

[10] Ossai, E. O., Nduka, U. C., Madukaife, M. S., Udom, A. U. & Ugwu, S. O. (2024). An extended Markov-switching model approach to latent heterogeneity in departmentalized manpower systems. Communications in Statistics-Theory and Methods, 53(19), 6957-6976. https://doi.org/10.1080/03610926.2023.2255322

[11] Guerry, M. A. & De Feyter, T. (2012). Optimal recruitment strategies in a multi-level manpower planning model. Journal of the Operational Research Society, 63(7), 931-940, 2012. https://doi.org/10.1057/jors.2011.99

[12] Dimitriou, V. A. & Georgiou, A. C. (2021). Introduction, analysis and asymptotic behavior of a multi-level manpower planning model in a continuous time setting under potential department contraction. Communications in Statistics-Theory and Methods, 50(5), 1173-1199. https://doi.org/10.1080/03610926.2019.1648827

[13] Guerry, M. A. & De Feyter, T. (2009). Markovian approaches in modeling workforce systems. Journal of Current Issues in Finance, Business and Economics, 2(4), 351-370.

[14] Dimitriou, V. A., Georgiou, A. C. & Tsantas, N. (2013). The multivariate non-homogeneous Markov manpower system in a departmental mobility framework. European Journal of Operational Research, 228(1), 112-121. doi.org/10.1016/j.ejor.2012.12.014

[15] Ahmed, M. A. (2017). The importance of the organizational structuring and departmentalization in workplace. The Journal of Middle East and North Africa Sciences, 3(3), 30-38. http://www.jomenas.org

[16] MacDonald, I. L. & Zucchini, W. (1997). Hidden Markov and other models for discrete-valued time series. CRC, 1997, 203.

[17] Bartholomew D. J, Forbes A. F. & McClean S. I. (1991). Statistical techniques for manpower planning (2nd ed). Chichester: Wiley.

[18] Neyman, J. & Pearson, E. S. (1933). The testing of statistical hypotheses in relation to probabilities a priori. Mathematical Proceedings of the Cambridge Philosophical Society, 29(4), 492-510.

[19] Perneger, T. V. (2021). How to use likelihood ratios to interpret evidence from randomized trials. Journal of Clinical Epidemiology, 136, 235-242. DOI: 10.1016/j.jclinepi.2021.04.010

[20] de Rochemonteix M., Napolioni V., Sanyal N., Belloy, M. E., Caporaso, N. E., Landi, M. T., et al. (2021). A likelihood ratio test for gene-environment interaction based on the trend effect of genotype under an additive risk model using the gene-environment independence assumption. American Journal of Epidemiology, 190(1), 129–141. https://doi.org/10.1093/aje/kwaa132

[21] Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. Biometrika, 60(2), 255-265.

[22] Bishop, C. M. (1995). Bayesian methods for neural networks. Microsoft, Report no. NCRG/95/009. http://www.ncrg.aston.ac.uk

[23] Helie, S. (2006). An introduction to model selection: tools and algorithms. Tutorials in Quantitative Methods for Psychology, 2(1), 1-10. 10.20982/tqmp.02.1.p001

[24] Ezugwu, V. O. & Ologun, S. (2017). Markov chain: a predictive model for manpower Planning. Journal of Applied Sciences and Environmental Management, 21(3), 557-565. https://dx.doi.org/10.4314/jasem.v21i3.17