DATA HANDLING - MODELING

Dick Snader: ... Thank you, Marty. First of all, I'm not doing this hour by myself. It's going to be a joint presentation, and I'd like to begin by introducing my colleagues. Immediately following me will be Owen M. Gleeson. Owen Gleeson is employed at U.S.F. & G. in the actuarial department. He's an Associate of the C.A.S. and very soon he'll be a Fellow. Following Owen, Bill Richards, from Aetna L & C will give a short talk. Bill is director of operations research in the corporate actuarial department.

We have many things to tell you about today, but unfortunately very little time has been allotted for the task. We therefore hope our presentation will whet your appetites and stimulate your interest enough to encourage you to correspond with us and to exchange ideas with us. I think, in the interest of conserving time, it would be best if you held your questions until all of us have made our presentations. We also have handouts for you to look at. The handout that I'm going to be following is captioned simply "Loss Reserve Symposium." I plan to stick very close to the text of the handout. So, you may wish to open it up and read along with me and, if you do, I would appreciate it if you would not read out loud. At the back of the handout are slides, copies of the slides that I'll be referring to. So, when they come on the screen, if you get tired of craning your neck to see the screen, you can just turn to the proper page in the handout.

The central problem of loss reserving is the problem of projecting ultimate claim costs. The problem has two major aspects. First, loss data must be properly organized. To accomplish this, a notational scheme is needed to facilitate the organization aid in the manipulation of data. Second, a statistically sound method of projection must be developed to operate on the data.

With respect to data organization, the two most common approaches are the report year method and the accident year method. The report year is usually defined by the date the claim records are created in the home office, and it is in this context that the term "report year" will be used in our presentation. It is convenient to say "report year" or "accident year", but the following discussion can refer to any fiscal period of any duration. When a report year approach such as the Fisher-Lange method is used, an attempt is made to measure the upward or downward development on claims which have already been reported to the company and to use this measure to estimate the aggregate reserve deficiency or redundancy on those claims. Now, report year approaches such as this may have many advantages, but no matter how good the method, I. B. N. R. cannot be obtained from it. To determine I. B. N. R., accident year analysis is required. One approach might be to use a report year method to test the adequacy of reserves on known claims and to use additional accident year data to estimate I. B. N. R. The accident year methods of this type are designed to measure only the emergence of I. B. N. R. All too often such plans fail to measure I. B. N. R. emergence beyond the first year following the reserve date.

Another approach, which is favored by many actuaries, is to use accident year data to measure the total development on all claims, both reported and unreported. This method has two drawbacks, one minor and one major. The first drawback is that the separation between I. B. N. R. and the reserve on reported claims is blurred and is usually defined arbitrarily. The major drawback, however, is that accident year methods usually fail to measure the full development on I. B. N. R. claims after they are reported.

Incidentally, before going on, I would like to take a moment to define some important terms. I think we will be able to speak more precisely if we confine the use of the word "development" to reported claims and use the word "emergence" when we speak of I. B. N. R. claims. This terminology is implicit in the remaining discussion. The real problem with both the accident year approach and the report year approach, as we know them, is that each is a two-dimensional attack to what I regard as a three-dimensional problem. Instead of choosing to organize information by report year or by accident year, we need to organize it both ways and arrange it into three-dimensional arrays, any element of which can be denoted by a triplet.

Stan, would you mind showing the first slide? Behold, a triplet. We hope you'll find this notational scheme pleasing, and perhaps consider using it in your own work. Using the notational scheme which we prefer, this triplet is L (i, j, k). "L" stand for incurred losses; "i" stands for accident year; "j" stands for report lag, which is the time elapsed between the occurrence and reporting of an accident, and "k" stands for the number of years of development subsequent to reporting. Such a three-dimensional array can be depicted geometrically, and for your edification, or perhaps your mystification, the next slide shows such a geometric representation. Now, in this representation, the initial observation of the accident year and the subsequent emergence of I. B. N. R. appears in the i, j plane. That's the horizontal plane. The development on reported reserves is depicted in the j, k plane. Only one accident year, 1971, is shown here. So, the accident year is represented by a series of points in three space. The accident year is being observed as of 12-31-75, and for the purpose of illustration, we are going to assume that an accident year is fully emerged after 5 years, and claims are fully developed 5 years after being reported. Thus, accident year 1971 is fully emerged, but only partially developed. Let's look at this just a little bit more carefully. L (71, 0, 0), which is the point closest to the origin, or at the origin, represents claims for accident year 1971, which have been reported in 1971, and evaluated as of 12-31-71. L (71, 0, 4), which is all the way out on the k axis represents claims for accident year 1971, which have been reported in 1971, and evaluated as of 12-31-75. Now, we might have a slight problem in semantics here. I said that an accident year will be considered to be developed fully after 5 years. Perhaps I should have said, "after 5 observations." You can see it on the slide, five observations of Accident Year 1971, Report Year 1971, the first one being at the origin, and the second one immediately above it, and five in total. L (71, 4, 0) represents claims for accident year 1971 which have been reported in 1975 and evaluated as of 12-31-75, and that's the point that's all the way out on the j axis. These are late reported claims. As a matter of fact, all of the points between the lines down at the bottom of the j axis are late reported claims. But, in order to be able to construct a data array such as this, three critical pieces of information should be in your data base. These items are accident date, report date and evaluation date. Now, the evaluation date is the most important of those, since the other dates are usually available. I don't mean that you necessarily have to record evaluation dates as data, but if you want to do something like this, your data processing system must be capable of preserving successive evaluations of the accident date, report date coordinates.

Could you show slide 3, Stan? All right, in this slide, accident year 1971 is shown as it would be represented at the end of 1979, when it will be fully emerged and fully developed. The dashed lines indicate development on the accident year, which is expected to occur between 12-31-75 and 12-31-79, and it's these values which must be estimated and carried as reserves as of 12-31-75 for claims that occurred in 1971. Could you show the next slide, Stan? In this next slide, two accident years are shown together. A little to the right of 1971 on the i axis appears accident year 1972. I'd like you to try to use your imagination for a moment and pretend that accident years '73, '74 and '75 have also been drawn into the picture. You can sort of imagine what they would look like. Accident year '72, incidentally, you can see, is only partially emerged, or is not fully emerged. If three more accident years were to be added to this diagram, and we were to look straight down onto the i, j. plane, we would see a diagram that somewhat resembles one constructed by Tom Fowler in his prize winning I. B. N. R. paper, which was published as a result of the 1971 Boleslaw Monic Foundation competition. That's slide number 5.

In this pictorial representation, the emergence of I. B. N. R. in the i, j, plane is depicted, but the subsequent development on those claims in the j, r, plane is neglected.

If we could just go back to slide 4 now? Using this schematic as a frame of reference, the reserving problem can be thought of as observing a large rectangular solid at rest on its base, and the base lies in the i, j, plane. Now, you have to use your imagination again. Pretend that accident years '73, '74 and '75 are shown on the drawing. Also pretend that accident years '70, '69 and so on, are on the drawing. Those accident years appear behind accident year 1971, and although they would be fully emerged with respect to I. B. N. R., they would only be partially developed. So, now you can sort of imagine that you're looking at part of a rectangular solid or a block. The block is composed of a number of little cubes, but some of the cubes are missing from one of the corners. If all of the cubes were present, we could measure the volume of the block and our liability would be known. Now, we can easily measure the volume of the incomplete solid, but our problem is to deduce from the shape that is in place what the volume of the final structure will be. Then, by subtraction, we can determine the volume of the cubes yet to be added, and that, I submit, is analagous to the loss reserving problem.

I would now like to touch very briefly on the second aspect of the overall problem. To make projections and forecasts, it is probably most convenient to return to the two-dimensional environment and employ two-dimensional projection techniques. This can be done safely as long as it is not forgotten that the problem is three-dimensional in nature. In our schematic analogy, we can first estimate the ultimate area of the base of the block and then estimate the ultimate height of the block in order to determine the ultimate volume. Each estimate is made separately and two-dimensionally, but neither estimate standing alone is sufficient to solve the problem. The two estimates must be combined. Well, similarly, it is necessary to measure in separate steps the emergence of I. B. N. R. claims and the development on reported claims. The separate steps can be handled two dimensionally by arranging data in the standard triangular or trapizoidal arrays, but the resulting projections must then be blended into a composite estimate of the total claim liability. Well, that ends my part of the presentation. I would now like to turn the discussion over to Owen Gleeson, who will develop the model a little further and show you how it can be used.

Owen Gleeson: Much of the investigative work on this model resulted from comparing two loss reserving systems which use the same data base. In the report year method, for purposes of this discussion, the system of reserving which:

A. Projects each report year to its k^{th} evaluation, following the close of the report year, and

B. Projects the emergence on each accident year through j periods following the close of the accident year and projects the yet to emerge claims to the k^{th} evaluation following their emergence.

We have the definition of the report year method at the top of the slide. The assumptions implicit in this definition are that:

1. Claim amounts are fully determined k years following the close of the year in which claims are reported, and

2. All claims which are to emerge from an accident year will emerge within j years following the close of the accident year.

The accident year method, for purposes of this discussion, is a system of reserving which projects the incurred loss for the accident year to its j^{th} evaluation following a close of the accident year. These descriptions can be expressed symbolically as follows: Assumed in these examples that j equals k equals 4, the symmetry here makes for a simpler formula. The evaluation of reserves as of 12-31-t, "t" being any calendar year. Then the current observation on the report year t minus m is the summation from j equals zero to 4 of this expression here, which again stands for incurred loss. Clearly each report year is comprised of j plus 1, and that would be 5 in this case, components. Projected ultimate evaluation of this report year is given by this summation here, so that the difference given by this third expression here is the development reserve on the report year. The total development reserve is obtained by summing over m, where m is greater than or equal to zero and less than or equal to 4; that is, we form this summation here. This then is the total development reserve on claims which have been reported. The next expression to be developed is that for IBNR, that is claims which have not yet emerged. The accident year t minus m has yielded the emergence of m components as of 12-31-t. The remaining components are given by this summation here, where the 4 indicates that these components are projected to the fourth evaluation following the emergence. The text has the third evaluation in it and it should be the fourth. The sum over n, where m is greater than or equal to zero and less than or equal to 4 given by this expression here, and this gives the total IBNR reserves then under the report year method. The total sum given by this last expression here is the total bulk reserve under the report year method. For the accident year method, we note that the current evaluation of accident year t minus m is the summation indicated here, the first summation on this slide. The ultimate projected value of this accident year, again according to the accident year method as defined here, is given by this summation here, so that the actuarial reserve for this accident year as of 12-31-t, is given by this expression, the third expression on the slide, which is simply the difference of the first two expressions. The sum over all m, where m is greater than or equal to zero and less than or equal to 4, is the actuarial reserve as of 12-31-t, and this would be given by this expression here. This is the total bulk reserve as derived under the accident year method, again as it's been defined here. Now, having derived the form of the bulk reserves under the two methods, I'd like to turn to a couple of applications that we've considered. First, as an application of this model, we examined the difference between the two reserves. A difference is formally written as in this first expression here, a rather lengthy expression. This can be simplified to this last expression here. I'd like to interpret the parts of this expression. The first double sum, that is this part here, is the future development on claims from the accident years k minus 4 minus 5 minus 6 and minus 7, which were reported in the last 4 years, that is, k, t minus one and two and three. The second double sum in the expression is the development on claims from the latests four accident years which will occur in the years following t plus 4. Estimates of these quantities may be used to assess the gaps in the accident year method with a given development period. Also, the quantities may be examined to compare results produced by a report year method with that produced by an accident year method. I referred to gaps a few minutes ago. Each one of these summations here, the first and the second, can be considered to be separate gaps in the accident year method. A second application of the model is the determination of the effects on calendar year results when an accident year method is used. It is again assumed here that the projections that are made are made accurately. The observations to date, on losses, are given by this first expression here. The bulk reserve under the accident year method again, is given by the second expression so that the total incurred to date is estimated by the accident year method as the sum of the above two expressions and is given in this third expression here. Now, this is as of 12-31-t. A similar expression gives the incurred loss to date as of 12-31-t plus one. Now, if we subtract the incurred as of 12-31-t from the incurred as of 12-31-t plus one, we wind up with this last expression on the slide here. The first summation in this expression is the expected quantity, which is the current estimate of incurred loss for the accident year, t plus one. The second summation is the development during the calendar year t plus one on claims from accident years t minus 4 and prior. Now, as a last application of this model, I'll make an assertion that's not difficult to demonstrate formally, and the assertion is that the accident year method with development period equal to j plus k that would be capital j plus capital k, produces the same bulk reserves, all other things being equal, as does a report year method with emergence period equal to j and a development period equal to k. I'll now turn the presentation over to Bill Richards. He will discuss some of the loss reserving techniques that he has developed in terms of this model. Thank you.

William Richards: ... Well, what I'm going to describe today is a working mechanized system that's patterned after the structure that Owen has just described. Before I do a short background as to what's happening at Aetna, we have two independent reserving systems in the company. The prime system, the one used to compute the reserves held, is maintained by the casualty division, and a secondary system, or monitoring system, is maintained by a corporate staff area reporting directly to the president. It is this monitoring system that I'm going to talk about today. It is a totally mechanized system. It's not an accident year method, but rather an accident month construction, but as Dick pointed out, the time periods don't matter. What I want to do is talk a little bit about how this system actually forecasts from your current point of development for a particular accident period to its ultimate, which basically hasn't been covered yet, and then a bit about the potential, for information you'd like to be extracted from this system structured in this manner. What I'm going to try and get across is that you can get a lot more information than simply the current reserve estimate. The monitoring system at Aetna starts out by once a month tapping into our claim data file and extracting a magnetic tape that contains historical information on incurred - well, number of claims and incurred dollars by accident month by emergence or development month. That tape is generally available to us on or about six working days after the closing of an accounting month, fed to us on a magnetic tape; we mount it and we run it from a time sharing terminal within our department. The lines that we carry on the tape currently are schedule P lines and those key divisions within those lines. For example, the personal lines, the personal accounts and the commercial areas and other breakdowns. We have the ability to call any line or any combination of lines with the program as it's set up by the data processing people, but currently we're only using schedule P. Now, this matrix is the representation I'm going to work from when I'm trying to describe the system. We have roughly in there right now about 164 accident months, and for each accident month for claims, the claim emergence by month and for dollars, the dollar development by month, and the structure is as such, and it's a cumulative pattern across in this direction, so that what we're looking at on the diagonal here is the current development as of whenever we receive the tape, in this particular case, 8-76, to demonstrate that I just did this a short time ago. Now, before we do anything else, we mount the tape and we feed this information through a series of statistical programs to try and iron out all the bugs that happen to be in the data. Basically, we're looking for two kinds of fluctuations in the data those movements that are peaks and valleys, if you will, that are repeated from year to year, in fact, are seasonal patterns. We identify those patterns and use those patterns in forecasting the future development. We're also looking for those patterns that appear to be random, show no cyclical repetition. Those patterns are removed from the data prior to forecasting the future. Once we've gotten through with the manipulations of the data base, then we begin our forecasting and the forecasting is - there's nothing fancy about it. It's basically the ratio method, which I suspect is familiar to everyone here. We'll take a particular accident month developed, for example as of the third report month, or whatever month you want to call it, and we'll look at how historically the pattern has moved from the third to the fourth report month, and based on that historical development, we will move this one month into the future in the forecast, and then we will continue that process, going one month at a time until we've forecasted the complete development of claims (I didn't point this out earlier, but I should) we do this process for claims and for average values, and only in the last step, where we multiply average values times claims, do we in fact get our estimate of the ultimate incurred dollars. All right, we move this pattern out as far as we can based on our data base, and roughly 13 years of history, we can only forecast out based on historical patterns approximately 13 years. For those lines such as for example "worker's compensations," where the pattern perhaps goes on - the development goes on beyond that period of time, there is an extrapalation routine, which analyses the developments up to the 13th year and if warranted extrapolates out to that point at which the development becomes flat. So hopefully what we've done is accounted for all the development of a particular accident period. Well, after we've gone through all this process, we come out to a point where we have the ultimate. So, this column of figures is the ultimate number of incurred claims and the ultimate average values for all prior accident periods. If I can step back just for a moment, this system in addition to forecasting the future, analyses its own track record in terms of the past. How has it been forecasting last month, the month before, and so on? If it identifies a consistent bias in its estimate; if it's consistently estimating 5 per cent high, 3 per cent low, it will take its own track record and correct its future forecast for that bias. If, in fact, the errors in the prior forecast appear to be random, no correction is made. Now, we have our estimate of ultimate. Unfortunately though, this is a mechanical process and in any mechanical process, basically a computer is dumb, and it's bound to make mistakes. So, we have to set some control somewhere along the way, and I want to show one control for claims and one control of sorts for average values. Now, remember, this column of ultimate claims, okay, what the system does is give me a graphical display of this, which I'm going to show in the next slide. After doing the forecasting it gives us a display, and here's time from January '63 to August '76, and now it displays the ultimate number of claims as estimated by the system this month over all time, seasonally adjusted. And what we're most interested in is how did the system handle the most recent accident months where the least amount of information is known and the greatest chance for error exists. So, I'll look at this or whoever is operating the system will actually look at this information and if, in fact, the last few accident periods are out of line with the trend of where we appear to be going, or the trend, plus what we perhaps know that the system doesn't know, but what we're doing in our business, what I can do if I want to, is override what the system has said, insert a different number into the machine and let it use my estimate rather than the system's. Generally, I won't do this override for more than perhaps the last two or three accident months. All right, after we put that new information in, now we have a final estimate of the ultimate number of claims. Now, for average values, this kind of a forecasting procedure is far more unstable for a longer period of time than it is for a number of claims, and so we have to do something a little different, and what I have designed as different in this particular system, again, to get back, we have time across the bottom and average values across the top, and on here are three different curves and they run together and I'm sure it's very difficult to see them, but there is one, which is the system estimate that I just got through describing with that matrix.

There is a second estimator built into the system, which in effect, is a forecasting equation derived using regression analysis and that is this line that I'm drawing here. The form of the equation is the average values equal Ae to the BX, where B is an annual growth and X is a time variable. So, now I have two estimates. I have an estimate from the trend line, and I'll refer to it as "trend line" from now on, and an estimate from the system, and then we have to decide how we're going to weight those two estimates, and the function that we built in here is to weight them based on the amount - how much that month has matured, that accident month. If it is a recent accident month, you will give more weight to the average value for example, the trend line. If it is a month that is mature, we'll give more weight to the system. Typically, if you're talking about a line like auto property damage, where the development pattern is fairly rapid, the trend line equation phases out very quickly. In a line like malpractice, where the trend is very, very slow, that trend line will carry weight for a longer period of time. Okay, the weighting of those two procedures then gives us an estimate of the average value. If I can just move back to this again, now we've got our final estimates of average values ultimately and our final estimate of numbers of claims ultimately, and multiply the two together, and now if you will, let's picture this matrix as if it's incurred dollars. Let's not talk about the two pieces. In order to compute the current reserve from hereon, it's just a matter of arithmetic. You have your ultimate incurred, minus the development to date, this diagonal, at each individual accident month, that is the actuarial reserve for that accident month, sum those differences across all accident months and you have the actuarial reserve, not the total reserve, but the actuarial reserve for that particular line. But you have much more than that, and that's what I want to get into next. If you take your ultimate as you see it today and subtract not the current diagonal, but the prior month's diagonal,in this display, you've restated last month's reserve, and if you move your diagonal back in time and constantly difference that from your ultimate column, you've restated all prior reserves. So, in addition to the day's reserve, you can get a brand new, fresh picture of what all prior reserves are. I want to make sure I got that point. One additional point, while I have this display up, is remember now the system is forecasting all slices in this matrix as it goes out. It's not moving from here to here, but rather moving step-by-step to that point. So, actually when the system was run last month, it didn't just do everything I described; it also actually estimated what it expected to see on that tape this month, and it remembers what it remembers. I don't want to make it seem like it's intelligent, but it compares what it estimated to see on that tape with what actually appeared this month in terms of numbers, claims and reported dollars, and within certain ranges, we say that is normal activity that has appeared on the tape. If, in fact, the amount of information that came on this month is abnormal, is outside of some expected range, that keys further investigation before we just plod ahead and produce the answers out of here. We go out and talk to people and find out what happened. Was there an error this month? Was there a processing day missing? What happened? If we discover that something happened that, in fact, may mislead this particular process, we can override some or all of the activity that happened in the current month.

We wouldn't override all, but - Basically what you would do, for example, if one processing day too many appeared this particular month (That's roughly 5 per cent of the activity.) factor out 5 per cent of the activity prior to doing the forecast. I don't use that option very often, but on occasion it has to be done.

All right, so much for the system. The output that comes out of this system ranges anywhere from one page - there's a variety of displays, and they range anywhere from one page in length to several hundred pages, and the displays can concentrate on reserve information or report activity information or accident period information, and that information can be summarized by months or by quarters or by years, and the information can show claim information, average value information or in total incurred dollar information, and more than that, there is a file that has the premium earned and premium written that's off on the side, which is fed into the system during the displaying phase so that we can see loss ratios, reserve to premium ratios and so on. The number of displays are - well, I don't want to oversell it, but virtually unlimited. Our favorite display is something like this. For each particular line, across the top, reserve dates; we can specify any reserve date we'd like to see. The first set of numbers is the "held". What was the held reserve at those various points in time? And then how is the system estimating those reserves at other points in time? For example, the 12-68 reserve valued by the monitoring system at 12-68, at 12-69, at 12-75, currently, which is at 8-76, in effect, display force track record of how the system is doing in terms of estimating these reserves. Probably the most interesting one is the current. We have our current estimate of what those reserves should be, using data right through, well 6 days ago, if you will, the 8-76 answers, and with simple arithmetic, you simply subtract the held from that current you've got an estimate in reserve adequacy at those points in time, a revalued estimate, and a little more arithmetic, take the change in that reserve adequacy and in effect you have an estimate of strength and weakening in those reserves as you see it today. Now, if we have a line on here that indicates published underwriting results that we had at various points in time and we factor out that strengthening or weakening, we have, in effect, given ourselves a complete new picture of what our underwriting trends have been, and in fact, they may be very different, I'm sure you're all aware, different that what actually gets published. So, this comes out every month. Now, I don't want to go into all the displays, but one last display, which deals with something that has absolutely nothing to do with reserving. I suspect that's wrong, but at least I suspect a lot of people think it has anything to do with reserving, and that is information that the rate makers need. There is information within the system that can be of some value to that. They have asked me to design a special routine to extract some of that information for them, and this is one of the displays that we give to the rate makers, nothing to do with our reservers that I know of. The accident year development factor - If you don't know what it is, I'm not going to try and explain it in too much detail, but they want to know the accident month development month factor developed at different points in time and they are probably not terribly interested in the 12-month factor or the 24-month factor, but rather in some factor like 15 months or the 18 months. So, what we do with this little routine is walk over to the terminal, mount the tape, call the routine that displays accident month development factors, specify a line we'd like to look at, a division within that line we'd like to look at, and what point in time we'd like to see the development, in this case, 18-month factor, and what it will produce is the current estimate of that accident month, 18-month factor, roughly 1.26 and a complete revalued historical pattern of how that factor has been moving over a time, so that in fact, you may not want to use this factor, but rather use trend, a forecast of where that factor's going based on where it's been going in the past. Okay fine, I've gone trhough all the displays, Just as a final comment, the structure that Owen and Dick and myself have been talking about, I hope I've got the point across that it's more than just computing the current reserve estimate, but rather giving you a wealth of information on what has happened in the past. And I guess that's about it.

Martin Bondy: One of the reactions I had, of course, was the tremendous amount of work that went into the development of these models and so forth, and one of the things that struck me and maybe it's because I'm not too bright, is that there is a lot in there which it's kind of difficult to communicate in a lecture environment, and I was just thinking, and we'd like to hear from you folks and maybe you'll put it on your evaluation forms the possibility of having on one topic like modeling a round table, maybe limited to 20 people or something like that, where there could be an exchange back and forth. Just think about that. Anyway, I think we're ready for some questions for these men, if any of you folks have any at this point - Let me start by asking a question then.

Question - Martin Bondy: Bill, you said something about if there seems to be a bias in your forecasts, and it's constant, you can correct for it by saying, "Well, if you're always 5 per cent low, just estimate it 5 per cent higher," or something like that. Do you go further and say, "This method ought to work. Why doesn't it work - Why is there a bias?" Do you ever make investigations like that, and if so, what sorts of things did you find out?

Answer - William Richards: The way the system is set up, as trends are changing from one - You have historical patterns that look like something and suddenly they're changing to something else. During that changing period the system is going to be wrong. It's going to start missing the forecast, and when it starts missing in a consistent manner, such as 5 per cent all the time, I may investigate to find out to determine that there is a changing pattern and once I've determined that it's changing, I say, okay, the system is working properly. The place where I do the investigation, however, is that if the error is random and large, then obviously the system is just not functioning with that particular line of business. Malpractice is a classic example. This is not the greatest system in the world for estimating the number of claims for malpractice, and in that case, where the error is large and random and not consistent in one direction or another, we do try and add something to the program or change features in the program that would try and counteract that observation.

Question - Sid Hammer (of the Home): The question really is not a question, but a statement, and directed at Bill Richards. In regard to the last point that you had made in terms of utilizing the reserve testing data base in procedure for evaluation of loss development using the ratemaking process, as an observation I had noted that and apparently you had recognized, that standard methods of utilizing loss development be reflected in ratemaking, when in the past or in the most immediate future reserve adjustments have been made, the standard loss development procedure will overstate the future development, and utilizing the reserve valuation method for evaluating loss development, valuating actually the loss reserves only, will avoid this bias. Is this one of the reasons or the origins of this outgrowth of the system?

Answer - William Richards: I'm afraid I'm not going to be able to answer that question in too much detail. We've given presentations of this system to many people in the casualty division. Some of those people were ratemakers, and those people came to me, and more or less described what they'd like to see, and all I did, in this particular case, was act as a programmer. It is simply to add a display that will give them what they're asking for, and I think I described accurately what I give them. What they do with it, I honestly don't know.

Question - Jack Smith (of Vokswagen): I direct the question at Mr. Richards. Does the system introduce or compare the variable of the speed of closure on the claims? In other words, are you comparing the per cent of pending or the per cent of closed with prior periods?

Answer - William Richards: The answer is no. There is nothing in that system that has paid or closed information. We expect - We are working on that, is what I guess is a classic answer, but it is one of my objectives for this year to add paid and closed information to that model, and at that point, I hope to handle the question that you're asking, but currently, no.

Question - Charles Foley (Geico): Mr. Richards, does your system attempt, or have you had any success with projecting variables that would change in the future that can't be seen in the historical base, such as inflation rates?

Answer - William Richards: It's almost as if I planned on that question. I recall now - I forgot to mention one key factor on those average value trend lines, where they came from. I didn't just dream them up. At 6-month intervals, we update those trend lines. We redo our regression analaysis, and during that analysis, which is external to the system; this is not in the system, we do take into account economic indicators. We subscribe to DRI. Perhaps a lot of you subscribe to other services. We analyze economic factors; we analyze changes in programs that the claim department is currently doing, changes in coverage, almost any factor that we can get our hands on in terms of talking to the people in the division. That may, in fact, change those growth patterns, and at that point, we will update those equations and hopefully the updated coefficients, the A and the B coefficient in effect is what gets fed into the system. Hopefully the new equation will keep up to date with what's going on in terms of the economy, the economic indicators. Yes, we do try and keep that in, but it's not built directly into the system. We do that externally at about 6-month intervals.

Question - Dick Wohl (Utica Mutual): ... and I guess this is to Mr. Richards also. Since your model is on a monthly basis, has any work been done to see whether the random noise in such a model doesn't almost overpower the information ability. In other words, if you did something maybe on a quarterly basis, or if you tried that, you might actually have an easier time analyzing the data. Has that been looked at? Since I know on your computer you can combine it any way you want.

Answer - William Richards: I don't think I understand the question. You mean, we may fluctuate on the month to month -

Question - Dick Wohl: Right. I've seen monthly models and the ones I've seen, it seems as if the fluctuations between months and the fluctuations in monthly things such as you mentioned, the number of working days, you spend so much time trying to go for that, that sometimes you don't end up with too much information, and I was wondering whether you had that kind of problem.

Answer - William Richards: I would say we have not had that problem. Perhaps it's because of the smoothing outer rhythm and the overrides that we have within the system that maybe that's what prevents us from doing wild fluctuations, but no, we haven't moved that wildly. It used to be a quarterly model and we were asked to convert it to a monthly system. I should point out though that when we forecast our estimate for a given month, we aren't saying that's the answer. We're saying that's the answer plus or minus whatever the track record of the system says the variability is. Perhaps plus or minus two per cent or three per cent, or maybe 10 per cent when you come to malpractice, but all we're interested in is a range. Are we within some range? And that range does not fluctuate all that wildly from month to month.

Question - Martin Bondy: Is there another further question from the floor? If not, I have another one for anybody. Bill, I think you said you made your estimates by division, which I assume is some sort of a geographical breakdown. Do you calculate your reserves on any basis other than overall company wide basis, as far as geography is concerned, and what sort of a basis do you use as far as product or line subdivision and why did you choose that?

Answer - William Richards: We calculate our reserves on an overall, country-wide basis.

Question - Martin Bondy: Do you think you might run into changing distributional distortions that way?

Answer - William Richards: It's possible. In our company, I think the state of our technology is, at the present time, countrywide is the best we can do. However, it seems that if we were to be concerned with estimating no fault reserves, if we aggregate all types of no fault claims together and try to draw conclusions about development patterns, we are probably masking development say with respect to a state like Michigan, as opposed to a state like Maryland. We're probably drawing incorrect conclusions with regard to any of those states, and probably overall also.

Question - Martin Bondy: Yes, right. You not only have a problem of knowing how well you're doing in Michigan versus Maryland, but if your volume or some base, and that's another question, is changing at a different rate in those two places, you may be hurting yourself overall.

Answer - William Richards: That's quite right. Well, in our company, that's a weakness in our current systems.

Question - Martin Bondy: One of the things I was going to do in my little introductory comments was to talk about some of the things we have thought about when we developed our I. B. N. R. system at Crum & Forster. One of them was, do you do it geographically, but we'll talk about that later. Another question I want to ask is what about the - Did I just see a question from the audience? Yes, Sir?

Question - Alan Richards (American Express): I'm not fully cognizant with all these things, but I'd like to ask Mr. Richards, if I understood him rightly, the model outputs some current information on actual activity within the company, and what I was curious about as to, if this is correct, understanding why you devoted part of the resources of the model to developing information that perhaps should be in the accounting or operational information systems? I thought I understood you to say that part of the output of the system, in a ddition to forecasts of the reserves and the development, was some current information on current activities within the company.

Answer - William Richards: I said the potential was there to do that. We do not - You're absolutely right. There are reports already in existence that probably are handled by data processing and sent to the proper people in terms of the number of claims that came in this month. That is not a major part of the output of our system. It's primarily for our use more than sent out to other people. We concentrate on those items that kick off of our forecasts ultimate incurreds. That's the information that's not available anywhere else, because a forecast came out of the system.

The central problem of loss reserving is the problem of projecting ultimate claims costs. The problem has two major aspects. First, loss data must be properly organized. To accomplish this, a notational scheme is needed to facilitate the organization and to aid in the manipulation of data. Second, a statistically sound method of projection must be developed to operate on the data.

With respect to data organization, the two most common approaches are the Report Year Method and the Accident Year Method. The Report Year is usually defined by the date that claim records are created in the home office, and it is in this context that the term "Report Year" will be used in our presentation. It is convenient to say Report Year or Accident Year, but the following discussion can refer to any fiscal period of any duration.

When a Report Year approach such as the Fisher-Lange Method is used, an attempt is made to measure the upward or downward development on claims which have already been reported to the company and to use this measure to estimate the aggregate reserve deficiency or redundancy on those claims. Report Year approaches may have many advantages; but no matter how good the method, IBNR cannot be obtained from it.

To determine IBNR, Accident Year analysis is required. One approach might be to use a Report Year Method to test the adequacy of reserves on known claims and to use additional accident year data to estimate IBNR. The Accident Year Methods of this type are designed to measure only the emergence of IBNR. All too often, such plans fail to measure IBNR emergence beyond the first year following the reserve date.

Another approach, which is favored by many actuaries, is to use Accident Year data to measure the total development on all claims, both reported and unreported. This method has two drawbacks; one minor and one major. The first drawback is that the separation between IBNR and the reserve on reported claims is blurred and usually defined arbitrarily. The major drawback is that Accident Year Methods usually fail to measure the full development on IBNR claims after they are reported.

The real problem with both the Accident Year approach and the Report Year approach is that each is a two dimensional attack on a three dimensional problem. Instead of organizing information by Report Year or by Accident Year, we need to organize it both ways and arrange it into three dimensional arrays any element of which can be denoted by a triplet.

Using the notational scheme which we prefer, this triplet is i,j,k. Associated with each i,j,k there is a number L(i,j,k), where L stands for incurred losses, i stands for accident year, j stands for report lag, which is the time elapsed between the occurrence and reporting of an accident, and k stands for the years of development subsequent to reporting. Such a three dimensional array can be depicted geometrically, and Exhibit I shows just such a geometric representation.

In this representation the initial observation of the accident year and the subsequent emergence of IBNR appears in the i, j plane. The development on reported reserves is depicted in the j, k plane. Only one accident year, 1971, is shown here. It is being observed as of 12/31/75, and for the purpose of illustration we are going to assume that an accident year is fully emerged at the end of the fourth calendar year following the accident year and claims are fully developed four years after being reported. Thus, accident year 1971 is fully emerged but only partially developed.

In order to be able to construct a data array such as this, three critical pieces of information should be in the data base. These items are the accident date, the report date and the evaluation date. The evaluation date is the most important of these since the other dates are usually always available. It does not necessarily have to be recorded, but your data processing system must be capable of preserving successive evaluations of the accident date, report date coordinates.

In Exhibit II accident year 1971 is shown as it would be represented at the end of 1979 when it will be fully emerged and fully developed. The dashed lines indicate development on the accident year which is expected to occur between 12/31/75 and 12/31/79. It is these values which must be estimated and carried as reserves as of 12/31/75.

In Exhibit III two accident years are shown together. If three more accident years were to be added and we were to look straight down on to the i, j plane, we would see a diagram that somewhat resembles one constructed by Tom Fowler in his prize-winning IBNR paper published as a result of the 1971 Boleslaw Monic Fund competition.

In this pictorial representation, which is illustrated in Exhibit IV, the emergence of IBNR in the i, j plane is depicted, but the subsequent development on those claims in the j, k plane is neglected.

Using the schematic depicted in Exhibit III as a frame of reference, an analogy can be constructed. The reserving problem can be thought of as observing a large rectangular solid at rest on its base. The base lies in the i, j plane. The block is composed of a number of little cubes, but some of the cubes are missing from one of the corners. If all of the cubes were present, we could measure the volume of the block and our liability would be known. We can easily measure the volume of the incomplete solid. But, our problem is to deduce from the shape that is in place what the volume of the final structure will be. Then, by subtraction we can determine the volume of the cubes yet to be added. In this analogy the portion of the solid that can be seen corresponds to values of L(i, j, k) which have been observed. This missing cubes correspond to values of L(i, j, k) that will be observed in the future.

I would now like to touch, very briefly, on the second aspect of the overall problem. To make projections and forecasts, it is probably most convenient to return to the two dimensional environment and employ two dimensional projection techniques. This can be done safely as long as it is not forgotten that the problem is three dimensional. In our schematic analogy, we can first estimate the ultimate area of the base of the block and then estimate the ultimate height of the block in order to determine the ultimate volume. Each estimate is made separately and two dimensionally. But, neither estimate standing alone is sufficient to solve the overall problem. The two estimates must be combined. Similarly, it is necessary to measure in separate steps the emergence of IBNR claims and the development on reported claims. The separate steps can be handled two dimensionally by arranging data in the standard triangular or trapezoidal arrays. The resulting projections must then be blended into a composite estimate of the total claim liability.

Much of the investigative work on this model resulted from comparing two loss reserving systems which used the same data base. The Report Year Method for purposes of this discussion is a system of reserving which

a. projects each Report Year to its K-th evaluation following the close of the Report Year.

b. projects the emergence on each Accident Year through J periods following the close of the Accident Year and projects the yet to emerge claims to the K-th evaluation following the emergence.

The assumptions implicit in the above definition are that:

1. claims amounts are fully determined K years following the close of the year in which the claims are reported.

2. all claims which are to emerge from an accident year will emerge within J years following the close of the accident year.

The Accident Year Method is a system of reserving which projects the incurred for the accident year to its J-th evaluation following the close of the accident year.

The above descriptions can be expressed symbolically as follows. (It is assumed in these examples that J = K = 4). The evaluation of reserves is "as of" 12/31/T . Then the current observation on report year T - n is

$$\sum_{j=0}^{j=4} L(T - n - j,\ j,\ n)\ .$$

Clearly each Report Year is comprised of J + 1 (5 in this case) components. The projected ultimate evaluation of this Report Year is

$$\sum_{j=0}^{j=4} L(T-n-j, j, 4)\ ,$$

so that the difference,

$$\sum_{j=0}^{j=4} \left[L(T-n-j,j,4) - L(T-n-j,j,n) \right],$$

is the "development" reserve for the report year. The total development reserve is obtained by summing over $0 \leq n < 4$. That is

$$\sum_{n=0}^{n=3} \sum_{j=0}^{j=4} \left[L(T-n-j,j,4) - L(T-n-j,j,n) \right].$$

The next expression to be developed is that for the IBNR. Accident year T - n has yielded the emergence of n components as of 12/31/T. The remaining components are

$$\sum_{j=n+1}^{j=4} L(T-n,j,4),$$

where "4" indicates that these components are projected to the 3rd evaluation following the emergence.

The sum over n, $0 \leq n < 4$, is

$$\sum_{n=0}^{n=3} \sum_{j=n+1}^{j=4} L(T-n,j,4).$$

The total sum

$$\sum_{n=0}^{n=3} \sum_{j=0}^{j=4} \left[L(T-n-j,j,4) - L(T-n-j,j,n) \right] + \sum_{n=0}^{n=3} \sum_{j=n+1}^{j=4} L(T-n,j,4),$$

is the total Report Year reserve.

For the Accident Year Method we note that the current evaluation of accident year T - n is

$$\sum_{j=o}^{j=n} L(T - n, j, n-j) \,.$$

The ultimate projected value for this accident year is

$$\sum_{j=0}^{j=4} L(T-n,j,4-j) \,,$$

so that the actuarial reserve for this accident year as of 12/31/T is

$$\sum_{j=0}^{j=4} L(T-n,j,4-j) - \sum_{j=o}^{j=n} L(T-n,j,n-j) \,.$$

The sum over all n where $0 \leqslant n < 4$ is the actuarial reserve as of 12/31/T

$$\sum_{n=0}^{n=3} \left\{ \sum_{j=0}^{j=4} L(T-n,j,4-j) - \sum_{j=o}^{j=n} L(T-n,j,n-j) \right\} .$$

First as an application of the model we will examine the difference between the two reserves. The difference is

$$\left\{ \sum_{n=0}^{n=3} \sum_{j=0}^{j=4} L(T-n-j,j,4) - L(T-n-j,j,n) + \sum_{n=0}^{n=3} \sum_{j=n+1}^{j=4} L(T-n,j,4) \right\}$$

$$- \left\{ \sum_{n=0}^{n=3} \left\{ \sum_{j=0}^{j=4} L(T-n,j,4-j) - \sum_{j=o}^{j=n} L(T-n,j,n-j) \right\} \right\} ,$$

which simplifies to

$$\sum_{n=4}^{n=7} \sum_{j=n-3}^{j=4} \Big[L(T-n,j,4) - L(T-n,j,n-j) \Big] + \sum_{n=0}^{n=3} \sum_{j=1}^{j=4} \Big[L(T-n,j,4) - L(T-n,j,4-j) \Big] .$$

The first double sum in this expression is the future development on claims from accident years T-4, T-5, T-6, T-7 which were reported within the last 4 years T, T-1, T-2, T-3.

The second double sum in the expression is the development on claims from the latest 4 accident years which will occur in the years following t = T+4.

Estimates of these quantities may be used to assess the "gaps" in an accident year method with a given development period.

The quantities may also be examined to compare results produced by a Report Year Method with that produced by an Accident Year Method.

A second application of the model is the determination of the effects on Calendar Year results when the Accident Year Method is used. It is again assumed here that the projections are made accurately.

The observations to date are

$$\sum_{n=0}^{\infty} \sum_{j=0}^{n} L(T-n,\ j,\ n-j) .$$

The bulk reserve is

$$\sum_{n=0}^{n=3} \left\{ \sum_{j=0}^{j=4} L(T-n,\ j,\ 4-j) - \sum_{j=0}^{j=n} L(T-n,\ j,\ n-j) \right\} .$$

Therefore, the total incurred to date, as estimated by the Accident Year Method is

$$\sum_{n=0}^{\infty} \sum_{j=0}^{n} L(T-n,\ j,\ n-j) + \sum_{n=0}^{n=3} \left\{ \sum_{j=0}^{j=4} L(T-n,\ j,\ 4-j) - \sum_{j=0}^{j=n} L(T-n,\ j,n-j) \right\}$$

This is as of 12/31/T.

A similar expression gives the incurred to date as of 12/31/T+1.

Subtracting the incurred as of 12/31/T from the incurred as of 12/31/T+1 yields

$$\sum_{j=0}^{j=4} L(T+1,j,4-j) + \sum_{n=5}^{\infty} \sum_{j=0}^{j=n-1} \left[L(T+1-n,j,n-j) - L(T+1-n,j,n-(j+1)) \right].$$

The first summation in this expression is the expected quantity which is the current estimate of incurred loss for accident year T+1.

The second summation is the development during calendar year T+1 on claims from accident years T-4 and prior.

The last application of this model which I will cite here today is in the form of an assertion which is not difficult to demonstrate formally.

The assertion is that an Accident Year Method with "development period" equal to J + K produces the same bulk reserves (all other things being equal) as does a Report Year Method with emergence period equal to J and development period equal to K.

Chicago

Owen M. Gleeson — September 17, 1976 — Richard H. Snader

$$L\,[\,i,\ j,\ k\,]$$

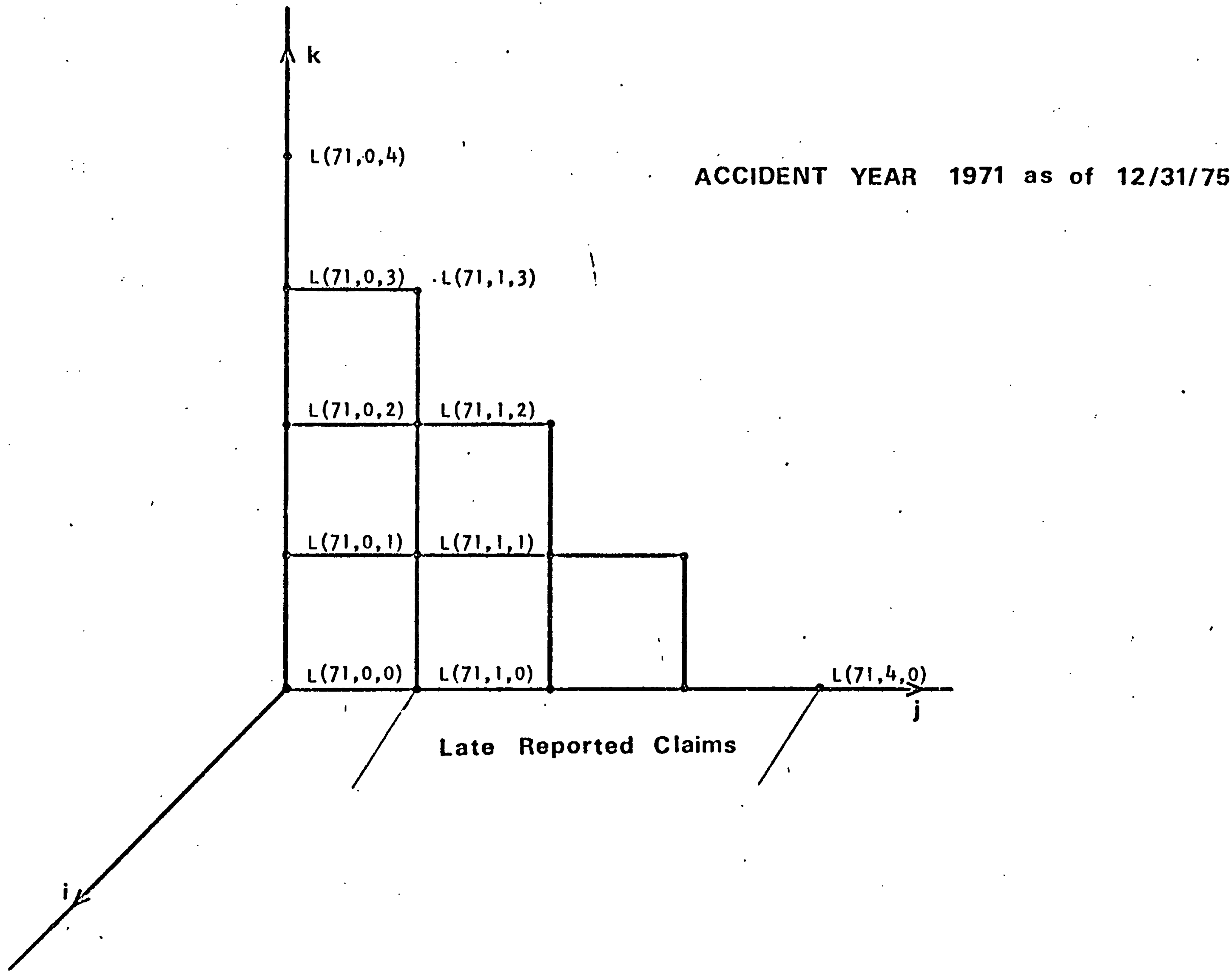
L = Losses

i = Accident Year

j = Report Year Lag

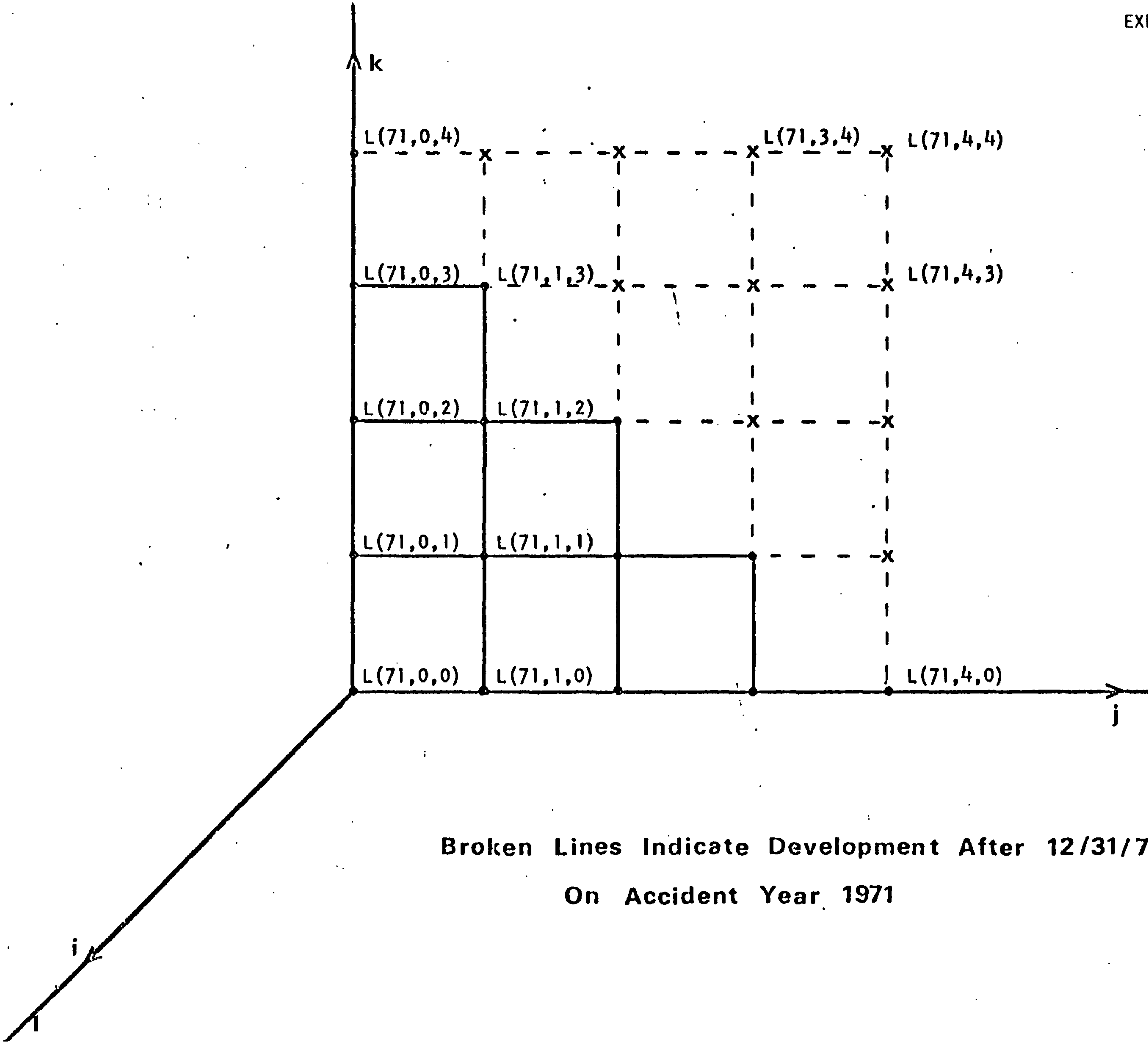
k = Development Year

EXHIBIT I



k
L(71,0,4)
ACCIDENT YEAR 1971 as of 12/31/75
L(71,0,3)
L(71,1,3)
L(71,0,2)
L(71,1,2)
L(71,0,1)
L(71,1,1)
L(71,0,0)
L(71,1,0)
L(71,4,0)
j
Late Reported Claims
i

EXHIBIT II



Broken Lines Indicate Development After 12/31/75

On Accident Year 1971

ACCIDENT YEARS 1971 and 1972 as of 12/31/75

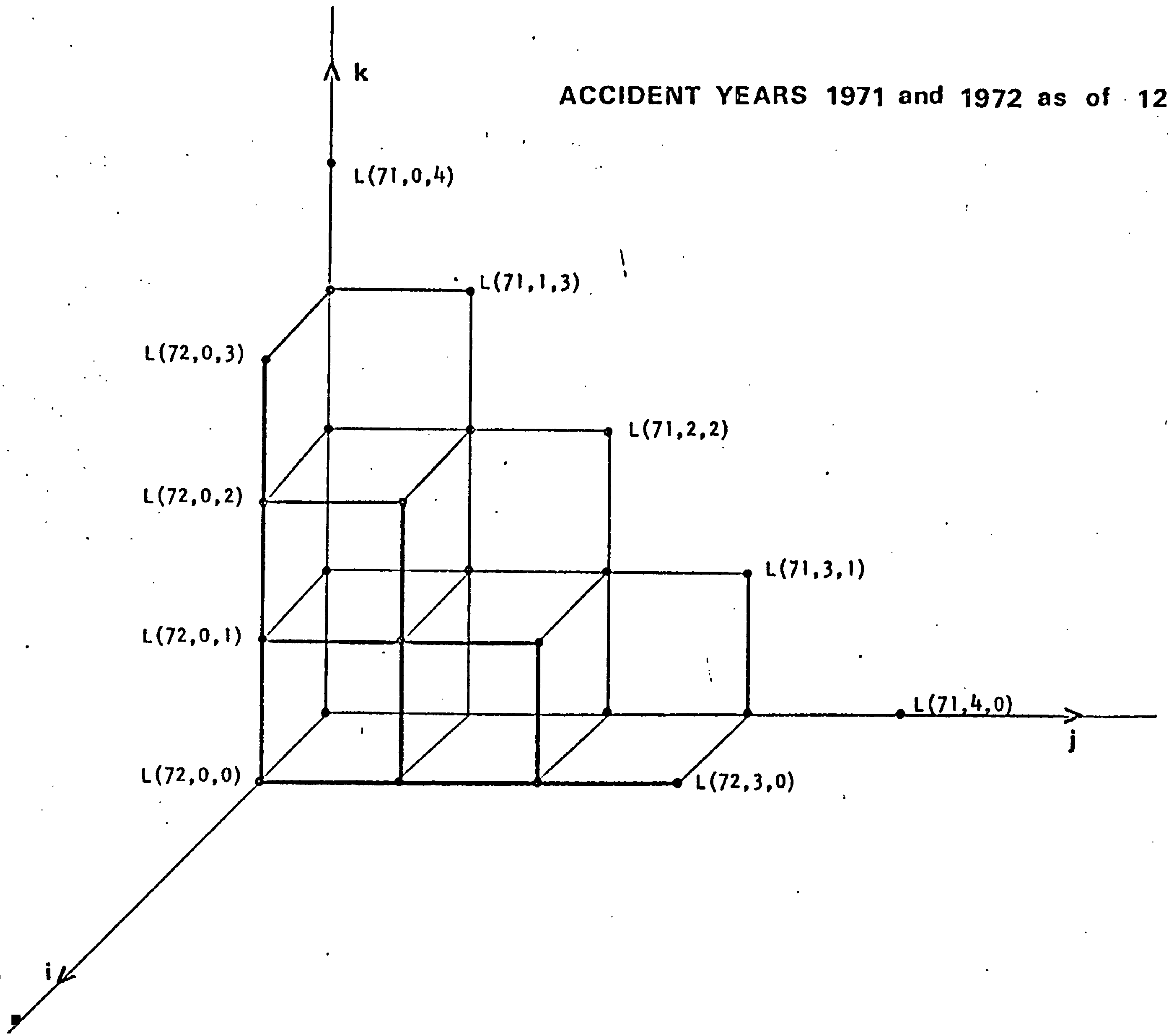
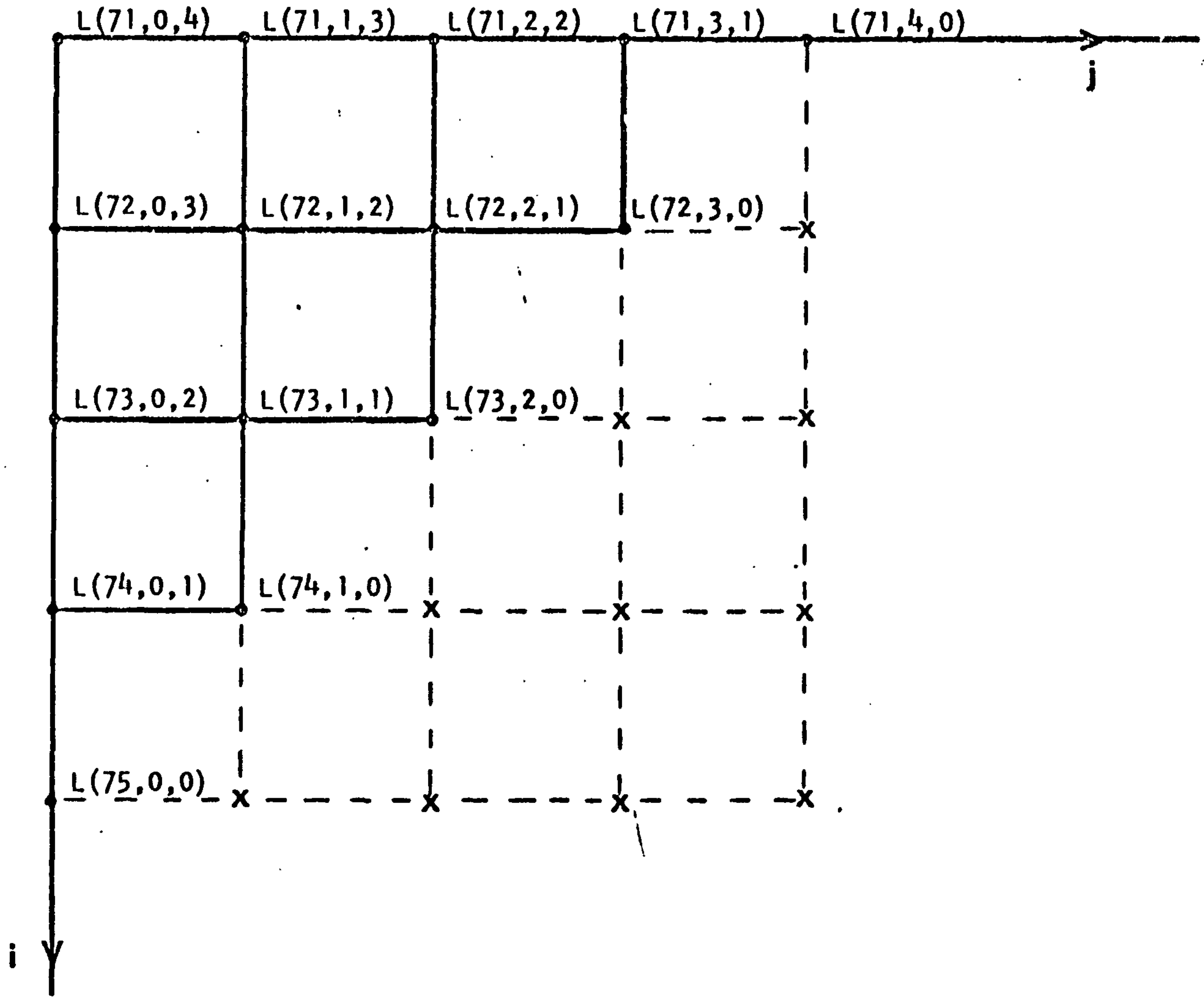


EXHIBIT IV


L(71,0,4)
L(71,1,3)
L(71,2,2)
L(71,3,1)
L(71,4,0)
j
L(72,0,3)
L(72,1,2)
L(72,2,1)
L(72,3,0)
L(73,0,2)
L(73,1,1)
L(73,2,0)
L(74,0,1)
L(74,1,0)
L(75,0,0)
i