Stochastic Farm Work Scheduling Algorithm Based on Short Range Weather Variation (Part 1)*
— Development of the Scheduling Algorithm —

I WAYAN ASTIKA*1, Akira SASAO*1, Kenshi SAKAI*1, Sakae SHIBUSAWA*1

Abstract

An algorithm for farm work scheduling named SFSW (Stochastic Farm Work Scheduling Algorithm Based on Short Range Weather Variation) was developed based on managing the risk of daily weather variation. The risk was evaluated by costs which were then minimized in the optimization. Genetic algorithms were utilized for optimization in order to attain flexibility.

The optimization was established through two steps; the first step was to decide the proper time to do jobs by minimizing the expected costs due to the effect of daily weather, and the second step was to decide the allocation of machinery by minimizing the costs of idle time of machinery and incomplete jobs. The schedule must be updated daily in order to extend the period covered by the schedule and adjust the schedule to the most recent conditions.

Keywords] farm work scheduling, genetic algorithms, optimum machinery allocation, short range weather forecast

I Introduction

Rain directly affects not only the quality but also quantity of operations in the field or indirectly affects other components of production systems as operations in the fields are forced to be done improperly. Direct effects mainly occur in farm works that are sensitive to rain, such as grain harvesting. The quality of the grain may decrease by being harvested in the rain because of its high moisture content. On the other hand, if the harvesting is postponed a timeliness loss will arise1. Rain may also affect machinery operations in wet fields. In such a condition, trafficability and workability of machinery become low2). Indirect effects happen to other components of production systems as the jobs in the field are done improperly. For example, when more harvesting is done early in order to avoid rainy days, the volume and duration of storage will increase, leading to greater losses3). On the other hand, if the harvesting in the field is postponed due to rainy days, the supply to mills will not meet the demand, which may lead to penalty costs.

Daily weather is a factor often neglected in planning long-term farm works or in systems where the effect of rain is not dominant1,4~7). If the scheduling is to optimize the arrangement of daily operations, which are sensitive to rain, then the weather should be taken into account. Several methods for handling the weather factor have been developed, but they are usually designed for a specific farm work problem

* Partly presented at 56th JSAM Annual Meeting (Mie University), April 1997, 57th JSAM Annual Meeting (Yamagata University), April 1998, and 3rd IFAC/CIGR Workshop on Artificial Intelligence in Agriculture (Makuhari), April 1998

*1 東京農工大学農学部（183-8509 東京都府中市幸町3-5-8 042-367-5761）JSAM Member, Faculty of Agriculture, Tokyo University of Agriculture and Technology, Fuchu-shi, Saiwai-cho 3-5-8, Tokyo 183-8509 Japan

Conventional methods for scheduling activities
adopt either an optimization\textsuperscript{1(4)5(7)} or heuristic approach\textsuperscript{6(12)(13)}. Optimization, such as mathematical programming, produces an optimal schedule but suffers from rigidity when applied to complex systems\textsuperscript{16}. On the other hand, heuristic approaches are practical but do not guarantee optimality. The new emerging genetic algorithms (GA) are usually classified as heuristic methods, but they also perform searching processes to seek the optimum value. With this merit, GAs have been intensively studied, developed, and applied to complex optimizations and scheduling purposes\textsuperscript{16-23}.

This research aims to develop an algorithm for general farm work scheduling with reduced risk to short range weather variations. The algorithm employs a GA in order to allow modifications to be made when it is applied in a specific scheduling problem. The algorithm is hereafter named SFSW (Stochastic Farm Work Scheduling Algorithm Based on Short Range Weather Variation)

\section*{II Scheduling Algorithm}

Fig. 1 shows the formulation of the effect of rain in SFSW. There are a number of field jobs \textit{i} with their own respective appropriate timings \textit{T}_{i}. Appropriate timings may be decided based on the season, growth stage of the crop such as the peak of maturity, or based on the demands of other components of a crop management system, such as the daily demand of processing unit or customers. In grain harvesting, the appropriate harvesting time is mainly determined by the peak of maturity because the crop has a short optimum harvesting period. In sugarcane harvesting, in which the crops have longer harvesting periods, the demands of other components of the system become the main factors in deciding \textit{T}_{i} that is, the daily demand of cane to be milled.

If rain is likely to fall on \textit{T}_{i}, the jobs will be performed in the rain or will be shifted to other bright days, early or late. Both choices may have adverse effects, both on the quality and quantity of the jobs or on other components of the system. In grain harvesting, as mentioned above, the grain will have a high moisture content if it is harvested in the rain. On the other hand, if it is harvested earlier, the grain will probably not mature enough. In addition, the grain should be stored before drying, which induces a loss. On the contrary, if it is harvested late, a timeliness loss arises.

In sugarcane harvesting, losses also arise if the cane is stored as a result of earlier harvesting. On the other hand, if the harvesting is done late, the demand of the milling unit can not be fullfilled.

If a planning to do the jobs on \textit{T}_{i} involves the risk of adverse effects, the timing of jobs should be rearranged. The scheduling is designed to optimally rearrange job timings from the original \textit{T}_{i}. The proper schedule consists of job timings having the smallest risk adverse effects: both effects on the quality and quantity of the jobs and the effects on other components of the system. In the scheduling algorithm, all adverse effects are treated as costs which are minimized in the optimization.

Fig. 2 shows the scheduling algorithm of SFSW. In order to reduce complexity, the
The scheduling algorithm uses two steps of optimization; the first step is to decide the job timings, and the second step is to decide the allocation of machinery. The daily working areas obtained from the first optimization become inputs for the second optimization, while the maximum coverage of machinery is a constraint for the first optimization.

Both optimizations use a GA, in which solutions are generated and then evaluated. In deciding jobs timings, the schedule is generated and then evaluated with all possible future weather occurrences. The result of the evaluation is the expected cost. The algorithm is repeated until the schedule with the smallest expected cost is found. In deciding the machinery allocation, the allocation of all machines is generated and then evaluated. The result of evaluation is the expected cost due to idle time of machinery and incomplete jobs.

Short range weather variations are taken from local weather forecasts consisting of the probability of rain. Detailed information, such as the expected amount of rainfall, can not be incorporated because it is available only for a very short period of time. Therefore, only two weather variations are considered: rain or bright. Even a very small amount of rainfall is considered as rain. When the scheduling algorithm is applied to a specific scheduling problem, the implementation of the jobs on a rainy day depends on the kind of the jobs. Rain-sensitive jobs are not done, while other jobs are done with implications that quality and quantity of the jobs decrease.

### III Optimization and Genetic Algorithms

#### 1. Deciding job timings

Consider a period of \( N \) days for which local weather forecasts are available and a set of \( M \) jobs to be done within the period. The schedule is expressed with \( x_{ij} \) representing the area of the field belonging to job \( i \) to be done on day \( j \).

\[
x = \begin{bmatrix}
    x_{11}, x_{12}, \ldots, x_{1N} \\
    x_{21}, x_{22}, \ldots, x_{2N} \\
    \vdots & \vdots \\
    x_{M1}, x_{M2}, \ldots, x_{MN}
\end{bmatrix}
\]

The performance of a schedule depends on the accuracy in allocating \( x_{ij} \) under the coming weather conditions. As shown in Fig. 2a, with the daily weather condition of either rain or bright, and the length of \( N \) days for scheduling, there are \( 2^N \) possible combinations of coming weather occurrences \( k \). Accordingly, there are \( 2^N \) possible values of cost \( C_k \). With all \( 2^N C_k \) and their respective probability of occurrence \( p(O)_k \), an expected cost \( E(C) \) is determined representing the performance of the schedule.

\[
E(C) = \sum_k p(O)_k C_k
\]

The occurrence probability of combination \( k \) is derived by multiplying all individual probabilities of weather occurrences \( w_{kj} \). The probability of \( w_{kj} \) having value "rain" is equal to the probability of rain on that day \( p(R)_j \) and the probability of having value "bright" is 1-\( p(R)_j \).

\[
p(O_k) = \Pi_j p(w_{kj})
\]
The cost $C_k$ is the sum of the costs of each individual job $i$ which depends on job timing $j$, weather occurrence on that day $w_{kj}$, job area $x_{ij}$, and the sensitivity of job $i$ to rain. The sensitivity of jobs to rain is represented with $c_{ilO}$ standing for a unit cost for doing job $i$ on $l$ days relative to $T_i$ under weather condition $o$. In grain harvesting, $c_{ilO}$ expresses the cost of the loss of quantity and quality of grain as it is harvested early or late and/or harvested on rainy days.

$$p(w_{kj}) = \begin{cases} p(R) & \text{if } w_{kj} = \text{rain} \\ 1 - p(R) & \text{if } w_{kj} = \text{bright} \end{cases} \quad (4)$$

In some scheduling problems, the calculation of cost $C_k$ may need a special procedure with variable $c_{ilO}$ values rather than constants. For example, again in grain harvesting, for stored grain, $c_{ilO}$ can be determined if the length of storage period is known. To perform this calculation, a special procedure simulating the process of handling the grain in the storage and drying unit is required.

Apart from the cost caused by the decrease of quality of the jobs or products, other kinds of cost probably arise for other components of the system. For example, if the drying unit receives insufficient amount of grain, the drying unit and workers will be idle, which can be considered as costs. The calculation of these costs also need a special procedure.

Such procedures may need IF-THEN logical paths or iterative loops. This can not be performed in conventional optimization methods that usually need pre-defined constant coefficients. Here, the merit of using a GA is obvious.

The genetic algorithms may be a standard GA\(^{(18)}\) or other suitable GA such as messy GA\(^{(22,25)}\). The schedule is expressed with a chromosome consisting of $M \times N$ genes in which the value of $x_{ij}$ is expressed by the $(i-1)N+j$th gene. Every time a new chromosome is generated or an old chromosome is modified with genetic operators, such as mutation or crossover, the constraints and heuristics are checked to confirm that the chromosome is valid. If the chromosome is invalid, the process is repeated until a valid chromosome is obtained.

The calculation of $E(C)$ becomes the fitness function of GA. As minimizing $E(C)$ is the objective of the optimization, the chromosome with lowest $E(C)$ over elapsed generations is saved as temporary solution. Every time the following generations find a better chromosome, temporary solution is replaced with this chromosome. Temporary solution is considered as optimum solution when there is no more better chromosome found.

2. Constraints and heuristics for deciding job timings

The factors relating to farm works vary widely over agricultural production systems. The farm work scheduling problem in a certain production system has its own constraints characterized by specific local practices or specific personal decisions of a manager. Such constraints and heuristics play an important role in arranging the schedule.

Typical constraints in farm work scheduling problems are as follows.

(1) All jobs $i$ should be completed during the working period.

The sum of daily jobs $x_{ij}$ during the working period should meet the total area $X_i$.

$$\forall i, \Sigma_j x_{ij} = X_i \quad (8)$$

(2) The capacity of machinery should be enough to cover all scheduled jobs every day.

Let $q_{rij}$ be the working time, in fractions of a day, of machine $r$ for doing job $i$ on day $j$
and \( E(s_{r,i}) \) be the expected field capacity. The areas covered by all machines should be more than the scheduled areas.

\[
\forall j \forall r \sum_i q_{rij} E(s_{r,i}) \geq x_{ij} \tag{9}
\]

The expected field capacity \( E(s_{r,i}) \) comes from the field capacity when a machine is used on a rainy day \( s_{r | o=rain} \) and the field capacity when the machine is used on a bright day \( s_{r | o=bright} \).

\[
E(s_{r,i}) = p(R) \cdot s_{r | o=rain} + (1 - p(R)) \cdot s_{r | o=bright} \tag{10}
\]

The value of \( q_{rij} \) becomes less than 1 if a machine is not used on a whole day or is used on two or more jobs in a day. In this model, a machine is assumed to be used only for one job in a day and the working time to be only two values: whether a whole day of work or no work at all, scoring 1 or 0 respectively for \( q_{rij} \).

\[
\forall j \forall r \sum_i q_{rij} \leq 1 \tag{11}
\]

The capacity of machinery is enough to cover all scheduled jobs in day \( j \) if there is at least one set of \( q_{rij} \) allocations that satisfies Eqs. (8), (9), (10), and (11). This check can be done by solving the system equations with the simplex method, multiple choice integer program\(^{20}\), or genetic algorithms as was done for optimizing the allocation of machinery in Chapter III.3.

(3) Job order

The jobs usually have to be done in a certain order that can not be changed. For example, plowing must be done before harrowing, manual sugarcane planting must be done on the same day just after furrowing in order to preserve soil moisture, etc. This order must be maintained in arranging the job timings.

3. Allocation of machinery

The allocation of machinery is performed with the objective of minimizing the costs of idle time of machinery and incomplete jobs. The cost of idle time includes payment to operators and the machinery during the idle time, and the costs of incomplete jobs include the penalty cost for being unable to fulfill the demand of other components of the system, for example, failing to provide enough cane for milling.

The machinery allocation is decided for each day independently referring to the area of jobs scheduled on the corresponding day \( x_{ij} \). Machinery allocation is expressed using \( f_{ri} \) for machine \( r \) in job \( i \). Index day \( j \) does not exist in the expression because the allocation is done for each day independently.

A machine is assumed to visit only one field in a day with the working time being either a whole day of work or no work at all, scoring respectively 1 or 0 for \( f_{ri} \).

\[
f_{ri} = \begin{cases} 0 & \text{if } x_{ij} = 0 \text{ or } x_{ij} > 0 \text{ but machine } r \text{ is not allocated to do job } i \\ 1 & \text{if } x_{ij} > 0 \text{ and machine } r \text{ is allocated to do job } i \end{cases} \tag{12}
\]

The expected area of job \( i \) that can be covered \( E(v)_i \) is estimated from the sum of the individual expected daily field capacity of all machines \( E(s_{r,i}) \).

\[
E(v)_i = \sum_r E(s_{r,i}) \tag{13}
\]

\[
E(s_{r,i}) = \begin{cases} 0 & \text{if } f_{ri} = 0 \\ p(R) \cdot s_{r | o=rain} + (1 - p(R)) \cdot s_{r | o=bright} & \text{if } f_{ri} = 1 \end{cases} \tag{14}
\]

The machines allocated to job \( i \) have expected idle time \( E(e)_i \) if \( E(v)_i \) is more than the scheduled area \( x_{ij} \). Here, the unit of idle time is fractions of a day.

\[
\forall i E(e)_i = \begin{cases} \frac{E(v)_i \cdot x_{ij}}{E(v)_i} & \text{if } f_{ri} = 1 \text{ and } E(v)_i > x_{ij} \\ 0 & \text{if } f_{ri} = 0 \text{ or } E(v)_i \leq x_{ij} \end{cases} \tag{15}
\]

Expected incomplete area \( E(u)_i \) exists if \( E(v)_i \) is less than the scheduled area \( x_{ij} \).

\[
E(u)_i = \begin{cases} x_{ij} - E(v)_i & \text{if } E(v)_i < x_{ij} \\ 0 & \text{if } E(v)_i \geq x_{ij} \end{cases} \tag{16}
\]

Let \( E(MC) \) be the expected total cost for idle time of machinery and incomplete jobs. Minimizing \( E(MC) \) is thus the function used to evaluate the optimization.

\[
E(MC) = \sum_r E(e)_r \cdot e + \sum_i E(u)_i \cdot u \tag{17}
\]

The machinery allocation is also decided through an optimization with a GA. The chro-
mosome consists of $Q$ genes in which the $r$th gene expresses the allocation of machine $r$. If machine $r$ is allocated to do job $i$, the value of the $r$th gene is set to $i$. For example, the 2nd gene having value 5 means that machine 2 is allocated to do job 5. In this case $f_{25}=1$. If machine 2 has been allocated to job 5, it can not be allocated to other jobs on the same day; $f_{2i}=0$. If a machine $r$ is not allocated to do any job, the value of the $r$th gene is set to 0.

From Eq. (12), if $x_{ij}=0$, or job $i$ is not done on that day, there is no need to allocate any machine to job $i$. In the chromosome, there will not be any gene with value $i$.

The fitness value of a chromosome is determined from Eqs. (13), (14), (15), (16), and (17). The constraints and heuristics must be checked. The most likely constraint is whether a certain machine $r$ can do a certain job $i$ or not.

4. Updating the schedule

The schedule optimized at the beginning is not suitable for the entire period because the weather forecast changes daily and works on a day may produce new conditions such as the remainder of jobs and machine troubles. In particular, weather forecasts tend to be accurate for one or two days ahead, but not for a week. In order to adjust the plan to the latest conditions and more accurate information, the schedule is updated daily as illustrated in Fig. 3.

IV Discussion and Further Improvements

The core of SFSW was configured by transforming the short range weather variations into costs due to changing job timings or decreasing job quality and quantity by rain. When the proposed model is applied to a scheduling problem, the procedure for determining these costs should be formulated. This is the most important part remaining to be developed.

Other parts of the algorithm also need to be changed based on the characteristics of the system. The most feasible parts to be changed are the constraints and heuristics. More complicated program coding will be required as the complexity of the system increases.

In deciding the jobs timings, the chromosome has $M \times N$ genes. These genes will be full if all $M$ jobs are done every day for $N$ days or all $x_{ij}>0$. This may be a very rare case rather than having many $x_{ij}=0$. In deciding the allocation of machinery, probably not all of the machinery need to be used; only one machine may be sufficient for all jobs. A messy GA, that has variable chromosome length in which only 'potential' genes are included, seems to be appropriate for this optimization. With a smaller number of genes, the search space can be reduced, and so the necessary computer memory and search time can also be reduced.

Also, weather condition is considered to be either rain or bright. For some farm work operations, a more specific consideration is probably required, such as the amount of daily rainfall. This algorithm could be improved by considering more specific weather conditions. One method that could be tried is to generate the amount of rainfall with a random number as was done by Dumont8). If the unit
costs $c_{i\ell o}$ for a specific amount of rainfall are available, the cost can be determined more accurately.

V Conclusion

The developed SFSW incorporates the weather variable by transforming the effects of weather variations into expected costs. The costs are then minimized in the optimization by using genetic algorithms. The scheduling process was divided into two steps: deciding the job timings and deciding the allocation of machinery.

References

13) Pitt, R. E.: A probability model for forage harvesting systems, Transaction of the ASAE, 25, 549-555, 1982

(Received: 3 Sep, 1998 - Question time limit: 31 May, 1999)

Notation

$\forall i$ : for all $i$
$C_k$ : cost when combination $k$ occurs
$c_{i\ell o}$ : unit cost when job $i$ is done on $l$ days relative to its appropriate day
$E(C)$ : expected cost
$E(e)_{r\ell}$ : expected idle time of machine $r$
$E(MC)$ : expected accumulation of cost due to idle time of machinery and incomplete jobs
$E(s_{r\ell})$ : expected field capacity of machine $r$
to do job $i$

$E(u)_i$ : expected area of scheduled job $i$ that can not be completed

$E(v)_i$ : expected area of job $i$ that can be covered

$e_r$ : unit cost for idle time of machine $r$

$f_{ri}$ : involvement of machine $r$ in doing job $i$; $f_{ri}=1$ if machine $r$ is allocated to do job $i$, and $f_{ri}=0$ if machine $r$ is not allocated to do job $i$

$M$ : number of jobs to be done within the scheduling period

$N$ : number of days within the scheduling period

$p(O)_k$ : occurrence probability of combination $k$

$p(R)_j$ : probability of rain on day $j$

$p(w_{kj})$ : occurrence probability of $w_{kj}$

$Q$ : number of machinery

$q_{rij}$ : working time of machine $r$ to do job $i$ on day $j$

$s_{rio}$ : capacity of machine $r$ to do job $i$ under weather condition $o$

$T_i$ : most appropriate day for performing job $i$

$u_i$ : unit penalty cost for incomplete job $i$

$w_{kj}$ : weather occurrence of day $j$ in combination $k$

$X_i$ : total area of field belonging to job $i$

$x_{ij}$ : area of field belonging to job $i$ to be done on day $j$

「研究論文」

天候の短期予測に基づく農業スケジューリングに関する研究（第1報）

—— スケジューリングアルゴリズムの作成 ——

イ. ウヤン、アスティカ*, 島尾 彰*, 酒井淳司*, 濱澤 栄*

要 旨

本研究は短期間（1週間）の天候予測に基づいて農作業のスケジュールを作成しようとするもので、天候予測のリスクをコストとして評価した。本報では作成したスケジューリングアルゴリズム、SFSW (Stochastic Farm Work Scheduling Algorithm Based on Short Range Weather Variation) について報告する。SFSWの中で、天候予測によるリスクをコストで評価し、これと最も似たとする最適化問題とした。この最適化には、遺伝的アルゴリズムを用いた。

最適化は二段階で行った。天候予測にしたがって、第一段階はコストを最小にする作業時間を決める。第二段階では雨天により作業できない時間と完了できない作業により予測されるコストを最小にし、また、作業する機械の最適配置を決定する。最新の天気予測を含むように、スケジュールは毎日更新されるようにした。

[キーワード] 農作業スケジューリング、遺伝的アルゴリズム、最適機械配置、短期天候予測