# Global Constraints for the Mean Absolute Deviation and the Variance: Application to the Vertical Line Balancing.

Pierre Schaus and Yves Deville
Department of Computing Science and Engineering
Université catholique de Louvain
2, Place Sainte-Barbe,B-1348 Louvain-la-Neuve (Belgique)
email: pierre.schaus@uclouvain.be and yves.deville@uclouvain.be

## Abstract

Optimization with a balancing objective often appear in practical problems where humans are implied in the solution. For example, in tasks assignment problems it is a desirable property that the workload is fairly distributed among the workers. In general, a violation measure of the perfect balance can be defined as the $L_p$ norm of the vector of variables minus their mean.

Two global constraints are presented that can be used in constraint programming to optimize the criteria $L_1$ (the mean absolute deviation) and $L_2$ (the variance).

These global constraints have been implemented in the Gecode and Ilog CP environments and experimented on the vertical assembly line balancing problem. To the best of our knowledge, this is the first exact method to minimize $L_1$ or $L_2$ on this problem.

**Keywords:** Constraint Programming, Balancing, Mean Absolute Deviation, Variance, Line Balancing.

Balancing constraints appears to be useful in many real applications especially when humans are implied. Indeed fairness among people is very important to produce acceptable assignments. Some examples of problems where the objective is to obtain well balanced solutions are: The Balanced Academic Curriculum Problem (BACP) [1], The (Vertical) Assembly Line Balancing Problem (ALBP) [3], Assigning patients to nurses [8], Balancing the assignment of customers among employees [6], Nurse rostering [13] and Generating Spatially Balanced Scientic Experiment Designs [4].

In more details, the ALBP is the following. A given number of workstations are placed along a conveyor belt. The workpieces are consecutively launched down the line from station to station until the end of the line. Some operations are performed on any workpieces in each station. The problem is to assign the operations to the workstations such that the workload of all the stations is nearly the same while satisfying various constraints such as precedences between operations [3].

In the ALBP, a hard balancing constraint would impose all the working stations to have the same workload $s/n$ where $n$ is the number of stations and $s$ is the total load (sum of load of operations). This often results in an over-constrained problem without solution. One possibility is to relax the hard balancing constraint with respect to some violation measure.

For a set of variables $\mathbf{X} = [X_1, X_2, ..., X_n]$ and a given fixed sum $s$, a violation measure of the perfect balance property can be defined as the $L_p$-norm of the vector $[\mathbf{X} - \mathbf{s/n}]$ with $\mathbf{s/n} = [s/n, s/n, ..., s/n]$ such that $\sum_{i=1}^{n} X_i = s$. The $L_p$-norm of $[\mathbf{X} - \mathbf{s/n}]$ is defined as $(\sum_{i=1}^{n} |X_i - s/n|^p)^{\frac{1}{p}}$ with $p \geq 0$.

Following the scheme proposed by Régin et al. [5] to soften global constraints, we define a violation of the perfect balance constraint as a cost variable $L_p$ in the global balance constraint. The constraint `soft-balance`$(\mathbf{X}, s, L_p)$ holds if and only if $L_p$-norm$([\mathbf{X} - \mathbf{s/n}]) = L_p$ and $\sum_{i=1}^{n} X_i = s$.

The interpretation of the violation for some specific norms is given below.

- $L_0$: $|\{X_i | i \in [1..n] \wedge X_i \neq s/n\}|$ is the number of values different from the mean.
- $L_1$: $\sum_{i \in [1..n]} |X_i - s/n|$ is the sum of deviations from the mean.

- $L_2$: $\sum_{i\in[1..n]}(X_i - s/n)^2$ is the sum of square deviations from the mean.
- $L_\infty$: $\max_{i\in[1..n]} |X_i - s/n|$ is the maximum deviation from the mean.

None of these balance criteria subsumes the others. For instance, the minimization of $L_1$ does not imply in general a minimization of criterion $L_2$.

Global constraints for criteria $L_1$ and $L_2$ are respectively `deviation` [11] and `spread` [9]. Given a sequence of finite domain integer variables $\mathbf{X} = [X_1, X_2, ..., X_n]$, a sum value $s$ and one variable $\Delta$, the constraint $\texttt{deviation}(\mathbf{X}, S, \Delta)$ holds if and only if $s = \sum_{i=1}^{n} X_i$ and $\sum_{i=1}^{n} |X_i - s/n| \leq \Delta$ and the constraint $\texttt{spread}(\mathbf{X}, S, \Delta)$ holds if and only if $s = \sum_{i=1}^{n} X_i$ and $\sum_{i=1}^{n}(X_i - s/n)^2 \leq \Delta$.

The $L_1$ criterion has already been applied on the vertical ALBP in [10] in a heuristic procedure and the $L_2$ criterion in [2] with genetics algorithms. We implemented `spread` and `deviation` in the Gecode and Ilog CP environments and applied those constraints on various problems. To the best of our knowledge, no exact method to optimize the balancing with respect to $L_1$ or $L_2$ has never been applied. The minimization of the maximum value is very popular to solve the ALBP [7]. Unfortunately this can result in very poor quality solutions from the point of view of $L_1$ or $L_2$.

The next Table illustrates this on the `Hahn` instance with 7 workstations from the benchmark data set of Scholl [12]. This instance has 53 tasks, the minimal duration of the tasks is 40 and the maximal is 1775. The results were obtained with Ilog Solver.

| constraint / measure | $\max_i\{X_i\}$ | $L_2$ | $L_1$ |
|---|---|---|---|
| time(sec) | 54.48 | **25.12** | 75.11 |
| $\max_i\{X_i\}$ | **2336** | (+3%)2400 | (+4%)2418 |
| $\sqrt{\sum_i(X_i - s/n)^2/n}$ | (+53%)1038 | **679** | (+9%)738 |
| $\sum_i \lvert X_i - s/n\rvert/n$ | (+34%)298 | (+0%) **222** | **222** |

It can be seen on the first column that, minimizing the maximum load among the stations can be a bad choice when the objective is to obtain nearly the same workloads for all the stations (vertical balancing). Indeed, the optimal solution obtained by minimizing the maximum workload is 53% suboptimal from best possible standard deviation and 34% suboptimal from the best mean absolute deviation. On this instance it seems preferable to optimize the $L_2$ criterion since the solution obtained is only 3% suboptimal for the maximum value criterion and is also optimal for the mean absolute deviation.

# References


[1] *Problem 30 of CSPLIB (www.csplib.org).*

[2] B. Rekiek et al. Applying the equal piles problem to balance assembly lines. *IEEE International Symposium on Assembly and Task Planning*, pages 399–404, 1999.

[3] Boysen et al. A classification of assembly line balancing problems. *European Journal of Operational Research*, pages 674–693, 2007.

[4] C. Gomes et al. The challenge of generating spatially balanced scientific experiment designs. *LNCS CP-AI-OR*, 2004.

[5] J.C. Régin et al. An original constraint based approach for solving over constrained problems. *LNCS CP2000*, 1894, 2000.

[6] Moutaz Khouja and Conrad Robert. Balancing the assignment of customer groups among employees. *International Journal of Operations and Production Management*, pages 76–85, 1995.

[7] Robert Klein and Armin Scholl. Maximizing the production rate in simple assembly line balancing: A branch and bound procedure. *European Journal of Operational Research*, 91(2):367–385, 1996.

[8] C Mullinax and M Lawley. Assigning patients to nurses in neonatal intensive care. *Journal of the Operational Research Society*, 53:25–35, 2002.

[9] G. Pesant and J.C. Régin. Spread: A balancing constraint based on statistics. *LNCS CP2005*, 3709:460–474, 2005.

[10] Rachamadugu R. and Talbot B. Improving the equality of workload assignments in assembly lines. *International Journal of Production Research*, 29:619–633, 1991.

[11] P. Schaus, Y. Deville, P. Dupont, and J.C. Régin. The deviation constraint. *LNCS CP-AI-OR*, 4510:269–284, 2007.

[12] Armin Scholl. Data of assembly line balancing problems. *Technische Universitt Darmstadt*, 93.

[13] Helmut Simonis. Models for global constraint applications. *Constraints*, 12:63–92, 2007.