Basketball team turn-over

The basketball team of the University of Bologna is not doing too well in the current season. The coach thinks that the problem is an imbalanced turn-over: the best players play too much and are often tired. In relatively easy games, the coach is forced to deploy the weaker players, to let the stronger ones rest. But, sometimes, this strategy leads to defeats even against weak teams.

We want to help the coach in having a more balanced turn-over and good chances of winning the next matches. Let $P = \{1, \ldots, p\}$ be the set of team's players. Let $F = \{1, \ldots, f\}$ the set of possible formations. Not every player can play in every position, and this means that not all 5-tuples of players are valid formations. For this reason, we explicitly list the feasible formations in set F. For convenience, let $\delta_{ij} \in \{0, 1\}$ be a parameter, which takes value 1 if player i is part of formation j, or 0 otherwise.

We want to come up with a strategy for the next K matches. The coach knows the strength and stamina of his players, and therefore gives us a table containing the number of minutes he would like each player i to play during match k. This information is encoded in a parameter m_{ik} which will have value 0 if the coach would prefer player i *not* to play during match k, and a value up to 40 otherwise (40 minutes being the duration of a match).

The parameters m_{ik} help us in building balanced strategies, but we also want to maximise the chance of winning the matches. A good indication of how strong two basketball teams are, relative to each other, is the "plus-minus score". Given a formation $j \in F$, the plus-minus score s_j tells us how many points, on average, were scored by our team when using formation j, compared to how many points were scored by the average team on our league.

For example, imagine that the average league team has scored 3.5 points per minute. This value can be calculated by dividing the total number of points scored by all teams during the season, by the number of teams in the league times the number of minutes played. When our team has used formation j, however, we have scored 3.8 points per minute. In this case, our plus-minus score would be $3.8 - 3.5 = 0.3$.

Analogously, given the team that we will meet for the k-th upcoming match, we can calculate their plus-minus score. This is done by dividing the toal number of points scored by that team during the season, by the number of teams in the league times the number of minutes played. This score will be indicated with S_k. Notice that, because we don't know which formation the team will use agains us, we are interested in calculating their overall *team* score, and not the score of any particular formation that they use.

With these data, we can build a model that maximises the chances of winning the nest K matches and, at the same time, minimises the excessive usage of the players, above the values m_{ik}.

Let $x_{jk} \in [0, 1]$ be a continuous variable, which represent for which fraction of match k we will use formation j. For example, if $x_{jk} = 0$, then we don't use formation j at all, during match k; if $x_{jk} = 1$ we let the players in formation j play for the whole 40 minutes of match k; if $x_{jk} = 0.5$, we use formation j for half of the match, and then we substitute at least 1 player in that formation.

Let $y_{ik}, x_{ik} \in \mathbb{N}$ two more variables, which represent the number of minutes that player i will play during match k. Variable y_{ik} will represent how many minutes more than m_{ik} player i will play during match k; if player i will not play any surplus minutes during match k, then y_{ik} will be 0. Variable z_{ik} will represent how many minutes fewer than m_{ik} player i will play during match k; again, if player i plays m_{ik} or more minutes during match k, then z_{ik} will be 0.

The reason why we use two separate variables ≥ 0, rather than one unconstrained variable, is that we want to penalise the situation when a player plays more minutes than the ideal value m_{ik}, but we don't want to penalise the situation when a player plays *fewer* minutes than m_{ik}, as long as we have good chances of winning the game.

Our model is, then, the following:

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{p} \sum_{k=1}^{K} y_{ik} && && (1)\\
\text{s.t.} \quad & \sum_{j=1}^{f} x_{jk} = 1 && \forall k \in K && (2)\\
& \sum_{j=1}^{f} s_j x_{jk} \geq S_k + 1 && \forall k \in K && (3)\\
& 40 \left(\sum_{j=1}^{f} \delta_{ij} x_{jk} \right) - y_{ik} + z_{ik} = m_{ik} && \forall i \in P,\ \forall k \in K && (4)\\
& x_{jk} \in [0, 1] && \forall j \in F,\ \forall k \in K && (5)\\
& y_{ik}, z_{ik} \in \mathbb{N} && \forall i \in P,\ \forall k \in K && (6)
\end{aligned}
$$

The objective function (1) minimises the number of surplus minutes played. Constraint (2) makes sure that, for each match k, we choose a list of formations that cover the whole duration of the match. Constraint (3) imposes that the expected number of points we score during the match is at least 1 more than the expected number of points that the opponents will score during the match. Finally, constraint (4) sets the variables y and z.

When using this model to create the strategies for the next games, however, the coach realises that the results are not always usable. Sometimes, for a match k, there are many variables x_{jk} which are > 0. For example, if there are 20 variables $x_{jk} > 0$ for a given k, that would correspond to 20 different formations to be used during one match, i.e. one change every 2 minutes. This is clearly not acceptable.

For this reason, the coach asks that at most 5 different strategies can be used during a match, and that each strategy cannot be used for less than 4 minutes (i.e. one tenth of a game). Can you fix the model, to comply with the coach's request? (*Hint:* the easier way might be to introduce a new binary variable w_{jk} that has value 1 iff trategy j is used during match k).