Traffic-Adaptive Base-Station Sleep Mode Control via Bayesian Optimization

Ryad Madi, Salah-Eddine Elayoubi, and Michel Kieffer
Université Paris-Saclay, CentraleSupélec, CNRS – L2S
F-91192 Gif-sur-Yvette, France
{ryad.madi, salaheddine.elayoubi, michel.kieffer}@centralesupelec.fr

***Abstract*—Base station (BS) sleep mode can reduce mobile-network energy consumption when traffic is low, provided that throughput and coverage requirements remain satisfied. The BS put in sleep mode must, however, be adapted as traffic changes, while evaluating candidate configurations through system-level simulations or direct network testing can be costly. We study a transfer-based Bayesian optimization (BO) approach for this traffic-adaptive BS deactivation problem. Random-Forest (RF) surrogates are trained once at a reference traffic demand and kept fixed online. Their throughput, coverage, and load predictions are adapted to the current demand through a low-order polynomial corrector fitted from accumulated observations. Corrected tree-level predictions are then used in a constrained acquisition function combining Expected Improvement and a learned feasibility probability. On a 21-BS system-level simulator with four traffic-demand levels, the proposed variants obtain lower energy consumption than a per-demand RF baseline and a contextual LinUCB baseline under matched evaluation budgets. An ablation study further shows that removing the traffic-dependent corrector degrades the final objective.**

***Index Terms*—Bayesian Optimization, Base-Station Sleep Mode, Network Energy Efficiency, Traffic-Adaptive Control.**

I. Introduction

Base stations (BSs) account for a substantial fraction of cellular-network energy consumption. As their power draw does not scale proportionally with carried traffic [1], a common energy-saving strategy is to deactivate selected BSs during low-load periods while trying to preserve Quality of Service (QoS) [2]. The deactivation decision is inherently traffic dependent: a configuration of BSs put in sleep mode that satisfies throughput and coverage constraints at low traffic may become infeasible when demand increases.

For a network with N BSs, binary sleep-mode control creates 2^N possible configurations. Costly simulation or network testing makes exhaustive search for the most energy-saving configuration satisfying the QoS constraints impractical. Bayesian optimization (BO) addresses expensive black-box optimization by using a surrogate model to select informative evaluations [3]. Here, the difficulty is that the operating condition itself changes: a surrogate learned considering a given traffic demand must be adapted before its predictions can guide decisions for another demand.

This work has been partially supported by the Sustainable 6G chair, funded by Orange and held by CentraleSupélec. It was also supported by the French government, in the framework of the France 2030 program (INTENTION-6G project).

We therefore study an application-specific transfer strategy for BS sleep-mode control. Random Forest (RF) surrogates are trained once for a reference demand and then are frozen. A small polynomial model is tuned to correct the RF outputs as online observations from other demands become available. The corrected RF predictions then drive a constrained acquisition function (AF). We evaluate this mechanism in both transfer directions (from the lowest and from the highest considered demand) and compare it with independent per-demand RF optimization and contextual LinUCB [4].

From an optimization perspective, the application combines three difficulties: a binary configuration space, QoS constraints, and an operating condition that changes between evaluations. Constrained BO commonly weights an AF by a probability of feasibility obtained from probabilistic surrogate models [5]. For discrete search, Sequential Model-based Algorithm Configuration (SMAC) uses RF surrogates and derives acquisition statistics from the ensemble [6], while d-MALIBOO combines BO with learned feasibility models for discrete constrained problems [7]. Contextual BO instead represents the context jointly with the decision variables, *e.g.*, through a GP over the context-action space [8]. Transfer BO methods reuse information from related source tasks through multi-task models or ensembles [9]. In contrast, the setting studied here starts from one RF model trained at a single traffic demand and transfers by correcting its outputs online.

II. Traffic-Dependent BS Sleep-Mode Selection

We consider N BSs serving mobile users over a given coverage area. A binary vector $\mathbf{x} \in \{0,1\}^N$ defines the network configuration, with $x_i = 1$ when BS i is active and $x_i = 0$ when in sleep mode. A traffic demand $d \in \mathcal{Z}$ is observed before the configuration is selected, with $\mathcal{Z}$ the set of traffic-demand levels. For each pair $(\mathbf{x}, d)$, a simulator returns cell-edge throughput $T(\mathbf{x}, d)$, defined for the worst 5% of users, coverage $C(\mathbf{x}, d)$, *i.e.*, the proportion of users whose SINR exceeds a target, and the per-BS load vector $\boldsymbol{\rho}(\mathbf{x}, d)$.

The BS sleep mode optimization problem is

$$\begin{aligned} \min_{\mathbf{x}\in\{0,1\}^N} \quad & f(\mathbf{x}, d) \\ \text{s.t.} \quad & T(\mathbf{x}, d) \geq T_{\min}, \quad C(\mathbf{x}, d) \geq C_{\min}, \end{aligned} \tag{1}$$

where, using the EARTH power model [1],

$$f(\mathbf{x}, d) = \sum_i x_i \big(P_0 + \Delta_P \rho_i(\mathbf{x}, d) P_{\max}\big). \tag{2}$$

Because the energy (2) requires the loads to be available, we predict $\boldsymbol{\rho}$ rather than fitting an additional objective surrogate.

III. Traffic-Adaptive BO

A. Frozen RFs and polynomial correction

At a reference demand d_0, separate RF surrogates are trained for throughput, coverage, and the load vector. RFs are well suited here because the decision vector is binary, inference over a large sampled candidate pool is inexpensive, and the ensemble provides per-tree predictions that can be reused as empirical uncertainty information [10]. Each RF contains $B = 10$ trees, uses a feature-subsampling ratio of $5/6$, and requires at least 10 samples per split. These RFs remain fixed throughout the online phase.

For an RF output $y \in \{T, C, \boldsymbol{\rho}\}$, let $\widehat{y}_0(\mathbf{x})$ denote its prediction at the reference demand d_0. At demand d, the prediction is adapted according to

$$\widehat{y}(\mathbf{x}, d) = \widehat{y}_0(\mathbf{x}) + r_{\boldsymbol{\theta}_y}(\widehat{y}_0(\mathbf{x}), \widetilde{d}), \tag{3}$$

where $\widetilde{d} = (d - d_0)/(\max \mathcal{Z} - \min \mathcal{Z})$ is the normalized traffic-demand offset.

Let $\mathcal{D} = \mathcal{D}_0 \cup \mathcal{D}_L$ denote the reference samples (in $\mathcal{D}_0$) together with L online observations accumulated across traffic levels. The reference samples anchor the correction at d_0, while the online samples characterize how the RF predictions deviate when the demand varies. For each output y, the polynomial-corrector parameters are obtained as

$$\widehat{\boldsymbol{\theta}}_y = \arg\min_{\boldsymbol{\theta}_y} \sum_{i \in \mathcal{D}} \left(\widehat{y}_0(\mathbf{x}_i) + r_{\boldsymbol{\theta}_y}(\widehat{y}_0(\mathbf{x}_i), \widetilde{d}_i) - y_i \right)^2 + \lambda_r \left\| \boldsymbol{\theta}_y \right\|^2, \tag{4}$$

The correction is learned from ensemble predictions and is subsequently applied to each individual tree prediction before computing the AF.

The correction is the identity for $L < 2$ online observations, linear for $L = 2$, and quadratic for $L \geq 3$. In the log-cost variants, throughput, load, and the resulting energy improvement are handled in log-space, while coverage remains linear.

B. Feasibility and acquisition

For each QoS constraint $g_j \in \{T_{\min} - T,\, C_{\min} - C\}$, the corrected RF ensemble gives the tree-satisfaction score

$$s_j^{\text{tree}}(\mathbf{x}, d) = \frac{1}{B} \sum_{b=1}^{B} \mathbb{1}[\widehat{g}_{j,b}(\mathbf{x}, d) \leq 0], \tag{5}$$

i.e., the fraction of trees predicting that constraint j is satisfied. A class-weighted logistic-regression classifier is refitted after each evaluation to estimate the joint feasibility probability $P_{\text{feas}}(\mathbf{x}, d)$. We consider a *direct* classifier using $(\mathbf{x}, d)$ and a *hybrid* classifier using

$$\boldsymbol{\phi} = [s_T^{\text{tree}}, s_C^{\text{tree}}, \mu_T, \sigma_T, \mu_C, \sigma_C, h(\mathbf{x})], \tag{6}$$

where μ and σ are the corrected (across-tree) means and standard deviations and $h(\mathbf{x})$ is the number of active BSs. Thus, the hybrid model compresses the N-dimensional configuration into RF-derived QoS statistics and one structural feature. Until both feasible and infeasible samples are available, the product $s_T^{\text{tree}} s_C^{\text{tree}}$ is used.

Corrected per-tree loads are substituted into (2), yielding $\widehat{f}_b(\mathbf{x}, d)$. With $f^\star(d)$ the best feasible energy consumption observed at demand d, we use the empirical tree-level Expected Improvement

$$\text{EI}(\mathbf{x}, d) = \frac{1}{B} \sum_{b=1}^{B} \max(0, f^\star(d) - \widehat{f}_b(\mathbf{x}, d)), \tag{7}$$

Rather than fitting a Gaussian approximation to the RF ensemble mean and variance as in SMAC [6], this criterion averages the improvement obtained directly from the corrected tree predictions. The same tree-level representation is therefore used on both the objective and feasibility sides. The constrained AF is [5]

$$\alpha(\mathbf{x}, d) = \text{EI}(\mathbf{x}, d) P_{\text{feas}}(\mathbf{x}, d). \tag{8}$$

We consider three variants: raw EI with hybrid feasibility (BO hyb.), log-EI with hybrid feasibility (BO log-hyb.), and log-EI with direct feasibility (BO log-dir.). At each online iteration, 10,000 candidate configurations with at least 10 active BSs are sampled directly in $\{0, 1\}^N$, the highest-scoring unevaluated candidate at the current demand is simulated, and the parameters of the corrector and of the feasibility classifier are updated. Direct discrete sampling avoids the relaxation-and-rounding mechanism used by some GP-based treatments of integer variables [11], [12], and the minimum-active-BS restriction excludes configurations known *a priori* to be outside the operating regime considered in the experiments. The RFs themselves are never retrained online. For the considered BS sleep mode selection problem, candidate scoring requires predictions of two QoS quantities and of the N load components across the B trees; with a fixed candidate-pool size and fixed RF surrogate, this is the dominant model-evaluation step. The online refits concern only the low-order ridge correctors and the logistic feasibility classifier on the accumulated data.

C. Online workflow

The key implementation choice in Algorithm 1 is to separate the expensive network load evaluation model from the traffic adaptation: RF inference always uses the fixed reference models, whereas only the low-dimensional correctors and the feasibility classifier are updated after a new observation provided by the simulator. A configuration already evaluated at the current demand is not selected again.

IV. Numerical Evaluation

A. Network and optimization setup

We simulate $N = 21$ BSs with average inter-site distance 450 m over 1.8 km^2, at 3.5 GHz with 90 MHz bandwidth, 4×8 BS antenna arrays, and $P_{\max} = 75$ W. Propagation uses path-loss exponent 3, 4 dB log-normal shadowing, and a 50-ray UMi multipath channel. User locations follow a homogeneous Poisson point process with user densities $\{30, 35, 40, 45\}$ users/km^2. A per-user flow-arrival rate of 0.1 and a mean file size 20 Mbits, yields traffic demands in the set

Algorithm 1: Traffic-adaptive BO for BS sleep control

Input: reference demand d_0, reference data $\mathcal{D}_0$, demand set $\mathcal{Z}$, online budget N_{eval}

Train RFs for T, C, and $\boldsymbol{\rho}$ at d_0; freeze them.
Initialize the correction data with $\mathcal{D}_0$.
for $t = 1$ **to** N_{eval} **do**
  Observe traffic demand d_t and sample candidate BS configurations.
  Obtain reference per-tree RF predictions for each candidate.
  Correct the predictions from d_0 to d_t using (3).
  Compute P_{feas} and, if a best feasible solution has been identified, EI.
  Select the highest-scoring configuration not yet evaluated at d_t.
  Evaluate it with the system-level simulator (or direct network testing).
  Add the observation to the shared data and refit the correctors and classifier.
  Update the best feasible configuration at d_t when improved.
end for

$\mathcal{Z} = \{60, 70, 80, 90\}$ Mb/s/km^2. The system-level simulator of [13] estimates T, C, and $\boldsymbol{\rho}$ over 100 independent user-location realizations and 500 channel realizations for each one. QoS thresholds are fixed from the fully active network at $d = 90$: $T_{\min} = 0.85 T_{\text{full}}$ and $C_{\min} = 0.95 C_{\text{full}}$.

Every method gets a budget of 40 offline and 40 online simulator evaluations. Results are averaged over five independent runs. The offline evaluation are done either at $d_0 = 60$ or at $d_0 = 90$ Mb/s/km^2: 10 feasible and 10 infeasible configurations are collected, followed by a 20-iteration BO refinement. The resulting RFs are then frozen. The two reference values are chosen to test the two extreme transfer directions available in $\mathcal{Z}$. With $d_0 = 60$ Mb/s/km^2, the fixed RFs are corrected toward higher loads. With $d_0 = 90$ Mb/s/km^2, they are corrected toward lower loads. Online demands are drawn uniformly from all four levels. For the two baselines, the 40 offline evaluations are spread as 10 configurations per demand. *Multi-RF* maintains and updates one independent RF per demand, with no transfer. *LinUCB* [4] uses one contextual linear model over configuration-demand features, with standardized energy as reward and penalties for throughput and coverage violations ($\lambda_T = \lambda_C = 10$), and exploration weight $c = 1$.

For LinUCB, the penalized reward used in the original comparison is

$$r_{\text{pen}}(\mathbf{x}, d) = -\frac{f(\mathbf{x}, d) - \mu_f^{\text{seed}}}{\sigma_f^{\text{seed}}} - \lambda_T \max\left(0, \frac{T_{\min} - T(\mathbf{x}, d)}{T_{\min}}\right) - \lambda_C \max\left(0, \frac{C_{\min} - C(\mathbf{x}, d)}{C_{\min}}\right), \quad (9)$$

TABLE I
MEAN ENERGY CONSUMPTION (W), WITH STANDARD DEVIATION OVER 5 INDEPENDENT RUNS, FOR THE TWO TESTED REFERENCE DEMANDS.

Method	$d = 60$	$d = 70$	$d = 80$	$d = 90$
Reference demand $d_0 = 60$				
BO hyb.	3226 ± 111	3435 ± 131	3558 ± 83	3965 ± 76
BO log-hyb.	$\mathbf{3163 \pm 109}$	$\mathbf{3342 \pm 14}$	$\mathbf{3525 \pm 63}$	3970 ± 123
BO log-dir.	3206 ± 75	3429 ± 71	3667 ± 64	$\mathbf{3939 \pm 70}$
Multi-RF	3434 ± 102	3639 ± 95	3962 ± 201	4203 ± 17
LinUCB	3366 ± 146	3511 ± 83	3781 ± 181	4133 ± 94
Reference demand $d_0 = 90$				
BO hyb.	3224 ± 108	3510 ± 86	$\mathbf{3573 \pm 65}$	4001 ± 47
BO log-hyb.	$\mathbf{3172 \pm 41}$	$\mathbf{3342 \pm 60}$	3595 ± 111	$\mathbf{3898 \pm 59}$
BO log-dir.	3184 ± 30	3430 ± 80	3691 ± 75	4046 ± 126
Multi-RF	3434 ± 102	3639 ± 95	3962 ± 201	4203 ± 17
LinUCB	3366 ± 146	3511 ± 83	3781 ± 181	4133 ± 94

where μ_f^{seed} and σ_f^{seed} are computed on the offline seed data. Demand is included in the LinUCB feature vector, so one linear model is shared across traffic levels. This baseline therefore tests whether a lightweight joint model of configuration and context can exploit the shared budget without the non-linear RF representation used by the proposed approach.

B. Energy results and adaptation speed

Table I reports the final energy at each demand for both reference choices. The proposed variants are below both baselines at every reported demand. BO log-hyb. has the lowest average across the four demands for both $d_0 = 60$ Mb/s/km^2 (3500 W) and $d_0 = 90$ Mb/s/km^2 (3501 W), compared with 3809 W for Multi-RF and 3698 W for LinUCB. Results are similar when transfer proceeds upward from 60 Mb/s/km^2 or downward from 90 Mb/s/km^2; this observation is limited to the two reference values tested here. No acquisition variant is best at every individual demand, although log-hyb. is best on average in both cases.

Fig. 1 shows the best feasible energy found so far across the evaluated contexts. The proposed variants improve rapidly during the first ten online evaluations and approach their final level after roughly 25. Multi-RF does not improve until about the ninth evaluation; because its online budget is split among four demand-specific RFs, each surrogate initially receives few new samples. LinUCB improves earlier but more slowly and with larger run-to-run variability. These curves complement Table I: Under the same online evaluation budget, the corrected reference model reaches lower best-feasible energy than the demand-specific Multi-RF baseline.

C. Corrector ablation and scope

Table II isolates the corrector for BO log-hyb. with $d_0 = 90$ Mb/s/km^2. With the RFs frozen but the correction disabled (Ablation A), the mean final objective increases from 3172 to 3360 W. A pooled-retraining strategy (Strategy B), which discards the frozen-reference design and retrains one RF after each evaluation using observations from all demands without demand as an input, reaches 3314 W. Thus, within this experimental protocol, explicit traffic-dependent correction performs

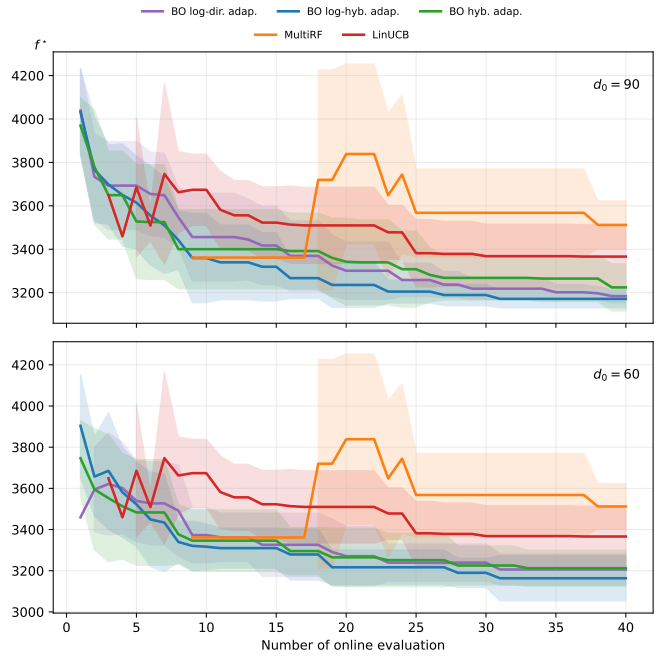


Fig. 1. Energy of the best feasible configuration $f^\star$ found so far across all evaluated contexts versus the number of online evaluations. Top: $d_0 = 90$ Mb/s/km^2. Bottom: $d_0 = 60$ Mb/s/km^2. Shaded areas show ± 1 standard deviation over 5 runs.

TABLE II
CORRECTOR ABLATION AND POOLED ONLINE RETRAINING FOR BO LOG-HYB., $d_0 = 90$ MB/S/KM^2; 5 RUNS.

Method	Mean final objective (W)
BO log-hyb. (proposed)	3172 ± 41
Ablation A: corrector removed	3360 ± 25
Strategy B: pooled RF retraining	3314 ± 54

better than either using the reference RF unchanged or pooling observations without modeling traffic.

D. Application-level interpretation and limitations

The comparison highlights the role of sharing information across traffic conditions. Multi-RF has the same surrogate family as the proposed method, but its data are partitioned across four independent models; during the early online iterations, each model receives only a fraction of the available observations. LinUCB shares all observations but imposes one linear representation over configuration and demand. The proposed approach occupies a middle ground in this experiment: it preserves the non-linear RF representation learned at one demand and uses the shared online data only to adapt the RF outputs. The ablation indicates that this explicit correction, rather than freezing the RF alone, is important for the observed result.

The frozen-RF design also changes what is updated online. Candidate scoring still requires RF inference over the sampled pool, but no RF fit is performed after the reference phase. Only the polynomial correctors and the logistic feasibility model are refitted on the accumulated data. Their refit cost does grow with the number of observations, so we do not claim that the total online computation is independent of the online history; the point is specifically that repeated RF retraining is avoided.

The conclusions are limited to the setting actually evaluated: binary BS sleep decisions, one scalar traffic-demand variable with four levels, two extreme reference demands, and simulator-based observations. The experiments do not determine the best reference-demand selection rule, do not evaluate more than four context levels, and do not include field or testbed measurements. The present study also does not establish convergence or correction-error guarantees. Vector-valued operating conditions, broader traffic scenarios, and theoretical guarantees are therefore left for separate work.

V. CONCLUSION

We applied transfer-based BO to traffic-adaptive BS sleep control. RF models of throughput, coverage, and BS load are trained once at a reference demand and then kept fixed, while a low-order polynomial corrector adapts their predictions online. On the considered 21-BS simulation scenario, the approach yields lower energy than per-demand RF optimization and contextual LinUCB under matched evaluation budgets for both tested transfer directions. The ablation confirms that the traffic-dependent correction is necessary to obtain the reported gain. The present results are empirical and simulation-based. Theoretical guarantees and broader context models are reserved for future work.

REFERENCES

[1] G. Auer *et al.*, "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, Oct. 2011.
[2] J. Wu, Y. Zhang, M. Zukerman, and E. K.-N. Yung, "Energy-efficient base-stations sleep-mode techniques in green cellular networks: A survey," *IEEE Commun. Surv. Tuts.*, vol. 17, no. 2, pp. 803–826, 2015.
[3] P. I. Frazier, "A tutorial on Bayesian optimization," arXiv:1807.02811, 2018.
[4] L. Li, W. Chu, J. Langford, and R. E. Schapire, "A contextual-bandit approach to personalized news article recommendation," in *Proc. WWW*, 2010, pp. 661–670.
[5] J. R. Gardner, M. J. Kusner, Z. E. Xu, K. Q. Weinberger, and J. P. Cunningham, "Bayesian optimization with inequality constraints," in *Proc. ICML*, 2014, pp. 937–945.
[6] F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Sequential model-based optimization for general algorithm configuration," in *Proc. LION 5*, 2011, pp. 507–523.
[7] R. Sala, B. Guindani, D. Ardagna, and A. Guglielmi, "d-MALIBOO: A Bayesian optimization framework for dealing with discrete variables," in *Proc. MASCOTS*, 2024, pp. 1–8.
[8] A. Krause and C. Ong, "Contextual Gaussian process bandit optimization," in *Proc. NIPS*, 2011, pp. 2447–2455.
[9] K. Swersky, J. Snoek, and R. P. Adams, "Multi-task Bayesian optimization," in *Proc. NIPS*, 2013, pp. 2004–2012.
[10] N. Meinshausen, "Quantile regression forests," *J. Mach. Learn. Res.*, vol. 7, pp. 983–999, 2006.
[11] P. Luong, S. K. Gupta, D. Nguyen, S. Rana, and S. Venkatesh, "Bayesian optimization with discrete variables," in *Proc. AJCAI*, 2019, pp. 473–484.
[12] E. C. Garrido-Merchán and D. Hernández-Lobato, "Dealing with categorical and integer-valued variables in Bayesian optimization with Gaussian processes," *Neurocomputing*, vol. 380, pp. 20–35, 2020.
[13] I. Merah, M. Kieffer, S.-E. Elayoubi, T. Clessienne, and B. Sayrac, "Towards network design-to-impact: RAN pruning for sustainability," in *Proc. IEEE ICC*, Glasgow, Scotland, 2026.