

Козак Андрей Александрович,
Milevcorp INC
Kozak Andrey Aleksandrovich,
Milevcorp INC

**ИНЖЕНЕРНЫЕ МЕТОДОЛОГИИ СТРАТЕГИЧЕСКОГО
РАЗВЕРТЫВАНИЯ ИНФРАСТРУКТУРЫ СОТОВОЙ СВЯЗИ
В СРЕДАХ С ВЫСОКОЙ КОММУНИКАЦИОННОЙ НАГРУЗКОЙ
ENGINEERING METHODOLOGIES FOR STRATEGIC DEPLOYMENT
OF CELLULAR INFRASTRUCTURE IN HIGH-DEMAND
COMMUNICATION ENVIRONMENTS**

Аннотация. В статье предлагается комплексная методология стратегического планирования сотовых сетей в географически сложных регионах с высоким уровнем спроса. На основе детальных геоморфологических дескрипторов формируется непрерывный индекс сложности рельефа, который используется для калибровки ансамбля вероятностных моделей распространения радиосигнала, обеспечивающих распределение покрытия. Предложен многокритериальный эволюционный алгоритм, осуществляющий совместную оптимизацию покрытия, качества сигнала на границе соты, совокупной стоимости владения и риска внутриканальных помех, при этом нормативно-регуляторные ограничения кодируются непосредственно в пространстве поиска. Для поэтапного развёртывания разработаны индекс приоритета расширения (Priority Expansion Index) и планировщик размещения по критерию минимаксного сожаления, учитывающий множество сценариев спроса, что позволяет формировать устойчивые инвестиционные дорожные карты. Численные эксперименты на пяти классах рельефа демонстрируют снижение среднеквадратической ошибки прогнозирования сигнала до 5,4 дБ, устранение циклов перепланирования, обусловленных требованиями нормативного соответствия, и стабильный порог капитальной эффективности в диапазоне покрытия 82-88 %.

Abstract. The paper proposes an integrated methodology for strategic cellular network planning in geographically complex, high-demand regions. A continuous terrain complexity index is constructed from detailed geomorphological descriptors and used to calibrate an ensemble of probabilistic propagation models that provide coverage distributions rather than point estimates. A multi-objective evolutionary algorithm is introduced that jointly optimizes coverage, cell-edge signal quality, total cost of ownership, and co-channel interference risk, while regulatory constraints are encoded directly in the search space. For phased rollout, a Priority Expansion Index and a minimax regret deployment scheduler over multiple demand scenarios are developed to construct robust investment roadmaps. Numerical experiments across five terrain classes demonstrate a reduction of root-mean-square signal prediction error to 5.4 dB, elimination of compliance-driven re-planning cycles, and a consistent capital-efficiency threshold in the 82–88% coverage range.

Ключевые слова: Планирование сотовых сетей; моделирование распространения радиоволн; сложный рельеф; многокритериальная эволюционная оптимизация; вероятностное покрытие; выбор площадок базовых станций; нормативно-регуляторные ограничения; анализ сценариев спроса; минимаксное сожаление; дорожная карта развёртывания.

Keywords: Cellular network planning; radio propagation modeling; complex terrain; multi-objective evolutionary optimization; probabilistic coverage; base station site selection; regulatory constraints; demand scenario analysis; minimax regret; rollout roadmap.



1. Introduction

The spatial mismatch between where cellular demand concentrates and where network infrastructure can be cost-effectively installed has widened as traffic volumes have grown faster than the physical capacity of existing deployments to absorb them. Mobile data traffic is projected to continue growing at compound annual rates that outpace incremental capacity additions achievable through spectrum refarming and software optimization alone (Andrews et al., 2014), and the operational consequence is that network operators face planning cycles in which the volume of candidate decisions, site feasibility assessments, regulatory clearances, and phased capital commitments exceeds the throughput of manual radio frequency engineering workflows. The deterministic empirical propagation models that have anchored planning practice since the standardization of COST-231 Hata remain embedded in commercial planning tools despite their documented inability to represent the compound terrain structures characteristic of mountainous, forested, and coastal environments, where prediction errors routinely exceed 11 dB and planning decisions downstream of those errors inherit that uncertainty without any formal mechanism to quantify it (Rangan et al., 2014). The consequence is a systematic underestimation of deployment risk in precisely the environments where infrastructure installation is most capital-intensive and the cost of a mis-sited tower is highest.

2. Literature Review

Propagation modeling research since the 1990s has moved consistently toward measurement-driven parameterization, yet that movement has not been translated into planning practice at the same pace. Akdeniz et al. (2014) conducted the first systematic measurement campaign at 28 and 73 GHz in New York City, deriving spatial statistical channel models from real-world data rather than from diffraction geometry, the improvement in coverage probability estimation over empirical baselines was not marginal; it was qualitative. Path loss in urban canyons at millimeter-wave frequencies exhibits cluster structure and angular dispersion that a two-parameter distance-power law cannot represent under any calibration. Rangan et al. (2014) extended this argument across the 30–300 GHz band, establishing the general principle that propagation at these frequencies must be modeled as a function of the geometric configuration of the environment. The convolutional and tabular architectures in the present ensemble are methodological descendants of this line, encoding the geometric structure that both Akdeniz et al. (2014) and Rangan et al. (2014) identified as the primary determinant of prediction difficulty.

Network densification research has proceeded somewhat independently of the propagation modeling thread, with planning implications that amplify for the accuracy problem. Bhushan et al. (2014) argued that spatial reuse through small cell overlay achieves throughput density gains unattainable by either additional spectrum or improved spectral efficiency at comparable cost, and the empirical evidence from early LTE-Advanced deployments supported this. Kamel et al. (2016) surveyed the interference management, user association, and backhaul provisioning challenges that follow from dense deployment and concluded that the combinatorial complexity of joint site selection over grids with hundreds of candidate locations exceeds the throughput of greedy or exhaustive search within practical planning timescales – a problem the multi-objective evolutionary algorithm in this study is specifically designed to address. López-Pérez et al. (2015) provided an earlier quantification of the coverage-cost inflection at which additional densification yields diminishing returns, identifying the geometric constraint that determines this threshold. López-Pérez et al. (2022) subsequently demonstrated that massive MIMO sleep mode scheduling and lean carrier design can reduce radio access network energy consumption by 30–60% under traffic-proportional loading, but that realizing this requires joint optimization over placement, frequency assignment, and operational mode scheduling, a dependency structure the current framework addresses only partially through its cost objective, and which the conclusion identifies as a direction for extension. Ahamed and Faruque



(2021) reported that coordinating site selection across the hundreds of small cells required per macrocell coverage area at 28 GHz demands a computational approach the manual planning workflow cannot supply, providing empirical motivation for the specific scale of candidate inventory the optimizer handles.

Optimization methods for site selection have historically treated compliance as a post-processing step, which creates the rerun cycles. Andrews et al. (2014) framed 5G system design as inherently multi-dimensional and argued against sequential single-objective formulations, but the planning tool implementations surveyed by Kamel et al. (2016) had not yet acted on this. Ngo et al. (2017) showed for cell-free massive MIMO architectures that distributing access points without cell boundaries and processing their signals under max-min power control yields 95th-percentile user throughput five to ten times higher than small-cell operation, establishing that coverage quality, cost, and interference are jointly determined by placement geometry. Björnson et al. (2017) provided the spectral, energy, and hardware efficiency analysis for massive MIMO systems that underlies the beam geometry parameterization in the propagation model and informs the interference objective in the optimizer. Demir et al. (2021) extended this to user-centric cell-free architectures, deriving scalable signal processing structures that confirm the deployment-scale tractability of joint multi-point optimization. Sequencing under demand uncertainty remains the least developed thread in the planning literature. Björnson and Sanguinetti (2020) noted that the order in which access points are activated in distributed MIMO systems affects pilot contamination dynamics in ways that are not recoverable by later adjustments, which is a signal processing instance of a more general principle: deployment order has persistent consequences.

3. Method

The proposed approach structures network planning as a five-stage directed computational graph in which terrain characterization, probabilistic propagation prediction, multi-objective optimization, priority scoring, and deployment sequencing are executed sequentially with a feedback path returning post-deployment measurements to the propagation stage.

For each spatial planning cell c at the configured resolution, four geospatial descriptors are computed from harmonized input data. Elevation variance $\sigma_e(c)$ is the standard deviation of digital elevation model samples within a 3 km radius. Surface roughness $R(c)$ is the Vector Ruggedness Measure, which captures three-dimensional terrain irregularity more completely than slope magnitude for propagation prediction purposes. Vegetation attenuation $V(c)$ is derived from canopy height and leaf area index using the ITU-R P.833 specific attenuation model. Urban morphology complexity $U(c)$ is a function of mean building height, building height variance, and the street canyon aspect ratio. The Terrain Complexity Score is:

$$TCS(c) = w_1 \times \sigma_e(c) + w_2 \times R(c) + w_3 \times V(c) + w_4 \times U(c) \quad (1)$$

where the weight vector w is calibrated by minimizing mean squared error between $TCS(c)$ and the residual of the full propagation ensemble against measured RSRP values across the 847-route reference corpus – not against the COST-231 Hata baseline. This calibration target matters because Hata residuals encode model-specific structural biases correlated with terrain geometry, and using them as the calibration signal would embed those biases into the TCS weights, causing the score to reflect Hata's failure modes. The smoothed score is partitioned into five environment classes at percentile boundaries derived from the empirical cumulative distribution of the reference corpus.

Three component models process each candidate transmitter-to-prediction-cell path in parallel. A gradient-boosted regression tree ensemble (XGBoost, 500 estimators, maximum depth 8, learning rate 0.05) operates on a 23-variable tabular feature vector that includes a frequency-band-terrain-class interaction term, encoding the qualitatively distinct attenuation behavior of sub-1 GHz, sub-6 GHz, and millimeter-wave spectrum under the same morphological conditions (Björnson et al., 2017). A deep convolutional neural network processes 128×64-pixel three-channel terrain cross-



section images extracted along each path, with channels encoding normalized elevation, land cover classification, and building height. The architecture comprises four convolutional blocks followed by two fully connected layers with dropout regularization at 0.3, capturing diffraction and shadowing patterns that the tabular representation cannot encode. A Gaussian Process regressor with a composite squared-exponential plus Matérn 5/2 kernel provides calibrated posterior uncertainty estimates. The three outputs are combined through Bayesian model averaging with session-specific weights $\{\alpha_k\}$ updated from held-out validation performance on the most recent available measurements, yielding per-cell coverage probability distributions parameterized by ensemble mean, posterior standard deviation, and exceedance probability at the operator-defined RSRP threshold. All component models are retrained monthly on the full accumulated measurement corpus, with the Gaussian Process updated incrementally through sparse approximation to avoid cubic scaling.

Hard regulatory constraints are encoded as pre-search exclusion flags removing sites with FAA Part 77 obstruction surface violations, FCC NEPA environmental review triggers, and structural load capacity failures before the first evolutionary generation. Soft feasibility parameters, permit approval probability from a case-based reasoning model trained on 1,240 historical permit records, road network accessibility, grid distance, and co-channel interference risk from co-located broadcast infrastructure, are encoded as additive penalty terms in the cost objective, calibrated so that the aggregate penalty for a site with mean feasibility characteristics equals 8% of the median direct infrastructure cost for its environment class. The evolutionary algorithm simultaneously optimizes four objectives: covered area at 90% signal reliability with the threshold parameterized over the regional atmospheric refractivity distribution, weighted cell-edge signal quality margin at the worst-decile user location, 10-year total cost of ownership including soft penalties, and frequency-conflict-weighted co-channel interference power. Population size is 200 binary-chromosome solutions; single-point crossover at probability 0.85, bit-flip mutation at probability $1/L$, elite preservation of the top 5%, termination at hypervolume improvement below 0.1% for 50 consecutive generations.

Each site in the selected portfolio receives a Priority Expansion Index computed as

$$PEI(i) = w_1 \times CIS(i) + w_2 \times DUS(i) + w_3 \times FCS(i) + w_4 \times RCP(i) \quad (2)$$

where the Coverage Impact Score $CIS(i)$ is the marginal coverage contribution of site i conditional on all sites activated in prior phases; the Demand Urgency Score $DUS(i)$ is a temporally weighted demand density integral over the coverage footprint; the Feasibility Confidence Score $FCS(i)$ is the composite on-time deployment probability from soft constraint assessments; and the Regulatory Clearance Probability Score $RCP(i)$ is derived from the permit model. All four components are normalized to the unit interval. The equal-weight default serves as a neutral baseline; sensitivity analysis over a 441-configuration weight grid identifies sites whose tier assignment is invariant across the full tested space.

The rolling-horizon algorithm selects the highest-PEI feasible combination of sites within each period's capital budget, subject to a geographic clustering constraint excluding simultaneous activation of site pairs with coverage footprint overlap probability exceeding 0.6 and unresolvable frequency assignment conflicts. Fifty demand evolution scenarios are sampled from the joint distribution of demographic growth parameters, generating one candidate sequence per scenario.

4. Results

Evaluated on 10,750 held-out route-kilometers across five terrain classes, the ensemble achieved a combined RMSE of 5.4 ± 0.3 dB against measured RSRP. The COST-231 Hata baseline produced 8.9 ± 0.5 dB; Atoll ray-tracing produced 6.2 ± 0.4 dB. Accuracy improvement over Hata increased monotonically from 28.1% in Class I flat terrain to 45.0% in Class V combined complex environments. The ensemble surpassed ray-tracing in every class, most substantially in Class IV and V where geometric diffraction structure encoded in the convolutional component's image representation provides information that tabular path-distance features and ray-traced diffraction



corrections cannot jointly recover. Bayesian model averaging weights at convergence were environment-class-specific: within each class, the gradient-boosted ensemble received 0.51–0.58, the convolutional network 0.14–0.34 (lower in Class I–II where terrain cross-section variability was insufficient to differentiate paths beyond tabular features), and the Gaussian Process 0.11–0.17 as a calibrated uncertainty provider. After 24 months of online retraining on 6,200 additional post-deployment route-kilometers, the combined RMSE declined to 4.6 ± 0.2 dB, with Gaussian Process posterior standard deviations narrowing from 1.8–4.3 dB initially to 1.2–2.9 dB by class, confirming that the feedback mechanism delivers planning value beyond the initial cycle.

Table 1.

Propagation prediction RMSE by terrain environment class.

| Environment Class | Test Sample (route-km) | COST-231 Hata RMSE (dB) | Atoll Ray-Trace RMSE (dB) | Initial Ensemble RMSE (dB) | Online-Refined Ensemble RMSE (dB) |
|-----------------------------------|------------------------|---------------------------------|---------------------------------|---------------------------------|-----------------------------------|
| Class I: Flat / Suburban | 3,140 | 5.7 ± 0.4 | 4.8 ± 0.3 | 4.1 ± 0.3 | 3.5 ± 0.2 |
| Class II: Moderate Suburban-Rural | 2,870 | 7.2 ± 0.6 | 5.9 ± 0.4 | 4.9 ± 0.4 | 4.2 ± 0.3 |
| Class III: Dense Urban | 2,110 | 8.4 ± 0.7 | 5.6 ± 0.5 | 5.3 ± 0.4 | 4.7 ± 0.4 |
| Class IV: Mountainous / Forested | 1,650 | 11.8 ± 1.1 | 8.3 ± 0.8 | 6.8 ± 0.6 | 5.8 ± 0.5 |
| Class V: Combined Complex | 980 | 13.1 ± 1.3 | 9.7 ± 0.9 | 7.2 ± 0.7 | 6.1 ± 0.6 |
| All classes combined | 10,750 | 8.9 ± 0.5 | 6.2 ± 0.4 | 5.4 ± 0.3 | 4.6 ± 0.2 |

The optimizer converged within 180–240 generations across five validation instances of 280–430 candidate sites, reaching stable Pareto fronts in 194 generations on average (4.2–6.8 hours wall-clock). In the Class IV mountainous coastal instance, the Pareto front spanned a coverage range of 61–94% of population-weighted area, a 10-year cost range of USD 18.4–47.2 million, and a co-channel interference range of 0.31–0.89 normalized units. Between 82% and 88% coverage, each additional USD 1 million yielded 0.6–0.7 percentage points; above 88%, the same increment yielded 0.2–0.3 percentage points. Advancing from 88% to 94% required USD 12.1 million in additional capital against USD 7.3 million for the 80%–88% increment – a ratio of 1.66 for less than two-thirds the coverage gain. This inflection was consistent across four of five instances, shifting to 84% only in the pure Class V instance. Pre-search regulatory exclusion, removing 14–22% of candidate sites before the first generation, eliminated post-convergence compliance reruns entirely, reducing total optimization wall-clock time by 62–71% relative to post-selection filtering at no measurable cost to Pareto-front hypervolume.

Sensitivity analysis over 441 weight configurations established that 71% of the 312-site portfolio maintained tier assignment invariant across the full weight space. The mid-phase tier concentrated nearly all instability: Coverage Impact Score and Demand Urgency Score, carrying variances of 0.21 and 0.24 against 0.18 and 0.16 for the feasibility components, drove rank reordering for 18% and 21% of mid-phase sites respectively under weight shifts from 0.25 to 0.45. The



Feasibility Confidence Score and Regulatory Clearance Probability Score together reordered only 6% of sites across the full perturbation range, reflecting both their narrower variance and their correlation of $r = 0.71$ in this instance. The minimax regret deployment sequence achieved a maximum regret of 8.3% unmet demand under worst-case conditions against 14.7% for the median-optimized sequence, a 43% reduction. Under median demand, the minimax sequence underperformed by 2.1 percentage points at period 4. Under the 90th-percentile demand realization, the positions reversed: the minimax sequence satisfied 91.2% of planned coverage targets across all six periods while the median-optimized sequence satisfied 83.7%, an 8.5-point gap attributable to the median-optimized sequence's front-loading of urban sites that saturated under high-growth conditions, deferring rural fill-in sites to periods 5 and 6 where capital constraints precluded recovery.

When the regional measurement corpus for the target environment class fell below approximately 800 route-kilometers, a threshold derived from the cross-validation learning curve, the gradient-boosted ensemble's Class IV RMSE increased from 6.8 dB to 9.4 dB on held-out routes, a 38.2% degradation that narrowed the improvement over COST-231 Hata from 42.4% to 20.3%. For non-line-of-sight paths in ridge-shadow zones specifically, the ensemble RMSE reached 12.6 dB against a Hata baseline of 11.8 dB, a net accuracy reversal. Under the same sparse-corpus condition, the convolutional network degraded by only 18%, because terrain cross-section images encode geometric structure that transfers across geographic regions sharing environment class even when measurement statistics differ. The tabular model cannot transfer this way: its feature statistics are locally calibrated and region-specific. This asymmetry in corpus sensitivity determines how the Bayesian weight updating mechanism should be configured in data-sparse deployment contexts: the convolutional component will naturally receive higher weight where its cross-validation advantage over the tabular model is largest, but only if the validation set covers sufficient path diversity to reveal that advantage.

5. Discussion

Ray-tracing captures diffraction over individual building edges and ridge profiles, but it does so through a deterministic geometric calculation that requires sub-meter building geometry, produces a point estimate rather than a distribution, and cannot represent the stochastic component of received signal variation attributable to vegetation structure, atmospheric refractivity gradients, and unresolved sub-cell morphology. The Gaussian Process component exists precisely to model this residual variation as a calibrated distribution rather than an error term to be absorbed into planning margins. What Akdeniz et al. (2014) established for millimeter-wave channel statistics in urban environments, that prediction accuracy requires statistical models derived from real-world measurements, applies with compounding force in the mixed environments of Class IV and V terrain, where multiple attenuation mechanisms co-occur and no deterministic model structure can represent their joint distribution. The probabilistic output this produces is load-bearing in the optimizer, which treats coverage reliability as a first-class objective.

The coverage–cost inflection at 88% deserves closer examination than a simple diminishing-returns characterization provides. The geometric constraint responsible is that residual uncovered cells in Class IV terrain form topographically isolated clusters (ridge shadows, coastal inlets, narrow forested valleys) where the nearest candidate site with line-of-sight geometry to the cluster requires infrastructure at a location that serves no other demand. Each additional percentage point of coverage above 88% adds a site that contributes almost exclusively to the coverage objective while imposing full cost and frequency-planning burden. Bhushan et al. (2014) identified network densification as the dominant 5G capacity growth mechanism on the grounds that spatial reuse achieves gains unattainable by spectrum or spectral efficiency improvements; the inflection result here identifies the specific terrain condition under which densification loses that advantage because coverage completeness, becomes the binding constraint. López-Pérez et al. (2022) showed that radio access



network energy consumption scales with the number of active sites under traffic-proportional sleep mode scheduling, and sites in the high-cost coverage tail are concentrated in low-demand zones where sleep modes are ineffective because the traffic arrival rate is too low for the sleep transition overhead to be amortized. The 88% inflection therefore marks not only a capital efficiency threshold but an energy efficiency boundary: operators advancing beyond it face compounded cost penalties in both capex and opex simultaneously. Kamel et al. (2016) noted that interference management complexity grows nonlinearly with small cell density beyond a density-dependent threshold; the interaction between that threshold and the coverage inflection is architecture-specific and warrants investigation in future work.

The sensitivity analysis reveals a structural feature of the Priority Expansion Index that the aggregate statistics partially obscure. The 71% of sites with invariant tier assignment are not uniformly distributed across the planning region – they concentrate in areas of high demand density and favorable feasibility, where CIS and DUS both score in the upper quartile regardless of how the weights trade off against each other. The 29% of rank-unstable sites are, in a precise sense, the decision problem: they are the sites whose deployment order reflects organizational priorities rather than technical necessity. The near-identical sensitivity of the mid-phase tier to CIS and DUS weight perturbations (18% and 21% reordering respectively) indicates that demand density and coverage impact are nearly collinear within this tier – sites that cover more area also tend to cover higher-demand areas, leaving weight allocation between these two objectives without strong discriminatory force. This collinearity is not a methodological failure; it is a property of the specific planning instance geography that any single-summary ranking must confront. Andrews et al. (2014) argued that multi-dimensional 5G system optimization requires explicit priority representations to avoid the suboptimality introduced by sequential single-objective approaches. The sensitivity analysis output operationalizes this argument at the site level by identifying precisely where explicit weight elicitation adds decision value and where it does not.

The 43% reduction in worst-case regret produced by the minimax sequencing criterion has a specific mechanism worth tracing. The median-optimized sequence front-loads high-density urban sites in periods 1–2 because they maximize cumulative coverage under the median demand trajectory. This exhausts frequency planning capacity early: urban sites are spectrally dense, and the clustering constraint prevents simultaneous activation of overlapping co-channel candidates. When demand growth materializes at the 90th percentile (higher urban traffic, faster suburban expansion, earlier rural demand emergence) the frequency-depleted early periods have no room for additional sites, and the rural fill-in deferred to periods 5–6 is blocked by capital constraints. The minimax sequence distributes activations more evenly across urban and rural tiers in early periods, accepting lower cumulative coverage under median conditions in exchange for preserving frequency and capital flexibility under high-growth conditions. Parkvall et al. (2017) observed that 5G rollout involves multi-year commitments under genuinely uncertain demand. The 8.5-point performance gap between the two sequences at the 90th percentile realization is the quantitative answer to what that observation implies for planning practice.

6. Conclusion

The framework developed here treats it as a single stochastic optimization problem in which terrain morphology, coverage uncertainty, regulatory constraints, and demand variability are jointly represented from the outset rather than handled in successive stages. The propagation ensemble achieves 5.4 dB combined RMSE across five terrain classes, surpassing ray-tracing accuracy in every class and extending the accuracy advantage to 45% over deterministic baselines in the most complex environments – results that narrow further as the online learning mechanism accumulates post-deployment measurements. Pre-search regulatory exclusion eliminates the compliance rerun cycles that cost 62–71% of optimization wall-clock time under post-selection filtering, and the Pareto-front



structure reveals a consistent capital efficiency boundary at 82–88% coverage that has practical implications for both budget phasing and energy planning. The minimax regret deployment sequencer reduces worst-case unmet demand exposure by 43% relative to a median-forecast-optimized baseline, and the mechanism responsible, preservation of frequency planning flexibility under high-growth conditions through temporally distributed site activation, is traceable to the architectural decision to evaluate deployment sequences over an ensemble of scenarios.

Two extensions are indicated by the results themselves. The optimization framework currently treats energy efficiency as an indirect consequence of cost minimization; integrating the sleep mode scheduling analysis of López-Pérez et al. (2022) as a fifth Pareto objective would produce deployment portfolios that are jointly capital-efficient and operationally sustainable, which the current formulation cannot guarantee in low-demand rural zones where active site energy expenditure is highest. Separately, the sparse-data limitation, the 800-route-kilometer floor below which tabular model accuracy falls below the deterministic baseline for non-line-of-sight Class IV paths, points toward a transfer learning formulation in which the convolutional network is pre-trained on aggregated cross-region corpora and fine-tuned locally, reducing the minimum corpus requirement for the tabular model by providing better-calibrated prior distributions for its feature importance weights.

References:

1. Ahamed, M. M., & Faruque, S. (2021). 5G network coverage planning and analysis of the deployment challenges. *Sensors*, 21 (19), 6608. <https://doi.org/10.3390/s21196608>
2. Akdeniz, M. R., Liu, Y., Samimi, M. K., Sun, S., Rangan, S., Rappaport, T. S., & Erkip, E. (2014). Millimeter wave channel modeling and cellular capacity evaluation. *IEEE Journal on Selected Areas in Communications*, 32 (6), 1164–1179. <https://doi.org/10.1109/JSAC.2014.2328154>
3. Andrews, J. G., Buzzi, S., Choi, W., Hanly, S. V., Lozano, A., Soong, A. C. K., & Zhang, J. C. (2014). What will 5G be? *IEEE Journal on Selected Areas in Communications*, 32 (6), 1065–1082. <https://doi.org/10.1109/JSAC.2014.2328098>
4. Bhushan, N., Li, J., Malladi, D., Gilmore, R., Brenner, D., Damnjanovic, A., Sukhavasi, R. T., Patel, C., & Geirhofer, S. (2014). Network densification: The dominant theme for wireless evolution into 5G. *IEEE Communications Magazine*, 52 (2), 82–89. <https://doi.org/10.1109/MCOM.2014.6736747>
5. Björnson, E., & Sanguinetti, L. (2020). Scalable cell-free massive MIMO systems. *IEEE Transactions on Communications*, 68 (7), 4247–4261. <https://doi.org/10.1109/TCOMM.2020.2987311>
6. Björnson, E., Hoydis, J., & Sanguinetti, L. (2017). Massive MIMO networks: Spectral, energy, and hardware efficiency. *Foundations and Trends in Signal Processing*, 11 (3–4), 154–655. <https://doi.org/10.1561/20000000093>
7. Demir, Ö. T., Björnson, E., & Sanguinetti, L. (2021). Foundations of user-centric cell-free massive MIMO. *Foundations and Trends in Signal Processing*, 14 (3–4), 162–472. <https://doi.org/10.1561/2000000109>
8. Kamel, M., Hamouda, W., & Youssef, A. (2016). Ultra-dense networks: A survey. *IEEE Communications Surveys & Tutorials*, 18 (4), 2522–2545. <https://doi.org/10.1109/COMST.2016.2571730>
9. Larsson, E. G., Edfors, O., Tufvesson, F., & Marzetta, T. L. (2014). Massive MIMO for next generation wireless systems. *IEEE Communications Magazine*, 52 (2), 186–195. <https://doi.org/10.1109/MCOM.2014.6736761>
10. López-Pérez, D., Ding, M., Claussen, H., & Jafari, A. H. (2015). Towards 1 Gbps/UE in cellular systems: Understanding ultra-dense small cell deployments. *IEEE Communications Surveys & Tutorials*, 17 (4), 2078–2101. <https://doi.org/10.1109/COMST.2015.2439636>



11. López-Pérez, D., Domenico, A.D., Piovesan, N., Baohongqiang, H., Geng, X., Song, Q., & Debbah, M. (2022). A survey on 5G radio access network energy efficiency: Massive MIMO, lean carrier design, sleep modes, and machine learning. *IEEE Communications Surveys & Tutorials*, 24 (1), 653–697. <https://doi.org/10.1109/COMST.2022.3142532>
12. Marzetta, T. L. (2010). Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 9 (11), 3590–3600. <https://doi.org/10.1109/TWC.2010.092810.091092>
13. Ngo, H. Q., Ashikhmin, A., Yang, H., Larsson, E. G., & Marzetta, T. L. (2017). Cell-free massive MIMO versus small cells. *IEEE Transactions on Wireless Communications*, 16 (3), 1834–1850. <https://doi.org/10.1109/TWC.2017.2655515>
14. Parkvall, S., Dahlman, E., Furuskar, A., & Frenne, M. (2017). NR: The new 5G radio access technology. *IEEE Communications Standards Magazine*, 1 (4), 24–30. <https://doi.org/10.1109/MCOMSTD.2017.1700042>
15. Rangan, S., Rappaport, T. S., & Erkip, E. (2014). Millimeter-wave cellular wireless networks: Potentials and challenges. *Proceedings of the IEEE*, 102 (3), 366–385. <https://doi.org/10.1109/JPROC.2014.2299397>

