

Urban Design Sustainability Through AI and Genetic Algorithms: San Felipe Case Study

Paulo César Shimabukuro Shimabukuro¹

¹ Lima, Perú
pshimabukuros@uni.com

Abstract. The research explores urban generative design in the San Felipe Residential Area using genetic algorithms, machine learning, and neural networks. Three urban scenarios are evaluated. The MEPS method (Spatial Metrics, Prediction and Segmentation) is introduced to analyze urban patterns and predict activities, producing an optimized pre-design whose urban spatial characteristics contribute to the sustainability of cities by maximizing their resources and minimizing their environmental impact.

Keywords: Machine Learning, Neural Networks, Genetic Algorithms, Architectural Optimization, MEPS Method

1 Introducción

La Residencial San Felipe (RSF) sirve como campo de estudio para registrar el impacto que tienen sus áreas recreativas y las actividades que se desarrollan.

Se propone una metodología basada en las Métricas Espaciales, Predicción y Segmentación (MEPS) que integra técnicas de aprendizaje automático, redes neuronales, algoritmos de optimización y diseño generativo; en un proceso de generar formas y espacios basados en optimizar los recursos como el área recreativa y la densidad desde un enfoque sostenible.

Se incorpora el criterio de “interacción” en el proceso de diseño, al predecir las actividades que pueden darse según algunas características del entorno; es una retroalimentación para la fase inicial del diseño que permite ir mejorando la propuesta.

El objetivo principal es validar el método MEPS, evaluando las “interacciones positivas” en tres escenarios generados por algoritmos.

Finalmente se aplicó la metodología al Conjunto Residencial Martinete como la contra parte a la RSF (La Cruz Minaya, 2022), tiene pocas “interacciones positivas”.

1.1 Revisión de la literatura

Actualmente se viene aplicando el aprendizaje automático en la investigación del diseño urbano y por mencionar algunas técnicas tenemos los Algoritmos Genéticos (GA), Redes Neuronales (Deep Learning), clustering (Machine Learning), diseño generativo optimizado (DPGO), mapas de calor (KDE), Geo-análisis con Métricas Espaciales, a continuación se presentan artículos y publicaciones que han contribuido en el proceso de investigación.

Enfocado en Shanghai, se investiga los parámetros de áreas recreativas para un diseño urbano sostenible en China (Wang & Mo, 2014), contribuyó como referencias para identificar patrones de uso respecto a sus tipos de entorno recreativo.

Analiza la expansión urbana, como la línea de tranvía en China, estudiando cómo afecta la dependencia de los automóviles (Zhou et al., 2019), modelo para identificar interacciones a través del Análisis Geoespacial, esto se aplica para determinar patrones en actividades distribuidas en la RSF. Para la visualización de los datos geoespaciales se usa la metodología para representar y analizar mapas generados por aprendizaje automático (Aghasi, 2018).

El Modelo Predictivo para evaluar los diseños entrenados con las MEPS mediante métodos de red neuronal, una referencia es la recuperación urbana post-COVID-19 (Ma et al., 2023), nos ayuda a comprender como crear y elegir redes neuronales adecuadas para la predicción.

Los métodos de segmentación o clustering pueden ser usados para comparar actividades humanas en entornos urbanos (Grauwin et al., 2014), esencial para interpretar el comportamiento agrupado y su relación con MEPS.

Diseño Optimizado Multi-Objetivo (diseño generativo y Algoritmos de optimización) se aplica por ejemplo para incrementar la eficiencia energética en aulas(Mo et al., 2022), nos sirve como referencia de como parametrizar en base al multi-objetivo de optimización. La Optimización mediante Algoritmos Genéticos son herramientas colaborativas en el diseño urbano (Yan et al., 2023), por ejemplo reducir el impacto que tiene urbanizar es un enfoque sostenible para la investigación. Aquí nos ayuda a encontrar el equilibrio entre la mayor área recreativa y la mayor cantidad de edificios.

Relaciones con el entorno se pueden cuantificar mediante Métricas Espaciales; en este caso usamos la librería “sf” de RStudio para manipular características geométricas (Pebesma, 2018), esta herramienta nos permite medir la proximidad, densidad e intersección con diferentes elementos del entorno.

Combinando la aplicación de estas técnicas se ha formulado el método MEPS y un análisis ACEP.

2 Metodología

2.1 Recolección de datos

Registro de observaciones continuas durante un mes desde las 7:00am hasta las 10:00pm por parte de un equipo de estudiantes de arquitectura distribuidos homogéneamente en el área de estudio. Cada individuo que realiza una actividad es registrado.

La recolección de datos está organizada en 15 variables de las cuales, por el enfoque experimental de investigación se decidió empezar por aquellas que ponen a prueba la metodología MEPS y el análisis ACEP.

2.2 Procesamiento y Análisis Exploratorio de los Datos (EDA)

En esta etapa, los datos recopilados son procesados y se verifican que el registro sea correcto impugnando las que presenten errores. Se realiza un análisis exploratorio para entender la naturaleza de los datos, detectar anomalías, y prepararlos para el análisis. Se utilizan técnicas como *One Hot Encoding* para convertir variables categóricas a variables binarias.

2.3 Extracción de Métricas Espaciales (EME)

Esta etapa consiste en la preparación de planos en QGIS que incluyen sendas, alamedas, edificios, parques-jardines y sombras proyectadas de árboles, que se consideran Elementos Espaciales Urbanos (EEU). Luego se importan a RStudio, donde se aplican comandos específicos para calcular distancias mínimas y densidades de EEU en un radio de 5 metros para cada punto de datos. Este proceso codifica la relación entre el entorno y la actividad realizada en valores numéricos conocidos como Métricas Espaciales (ME).

Tabla 1. Cuadro de variables

Datos de campo	Tipo de variable	Valores
Tipo de actividad	Categórica	0=Paseo-Caminando; 1 =Jugando-ejercicios, 3=Conversando; 5=Cuidando Niños; 8=Descansando
Entorno donde se realiza	Categórica	1=Grass; 2=Cemento; 3=Banca; 4=Polivalentes (usado de forma diferente para la cual fue diseñado)
Superficie requerida	Discreta	Metros cuadrados a partir de una estimación visual, por parte del observador capacitado
Coordenadas X, Y	Continua	Se registra en campo, luego se indexa a un plano CAD con coordenadas X, Y

Fuente: Elaboración propia, 2023

2.4 Análisis de Comportamiento, Entorno y Proximidad (ACEP)

Aquí se utilizan técnicas de segmentación o *clustering*, como el método 'K-means', y visualización con mapas de calor mediante la Estimación de la Densidad de Kernel (KDE). Se analizan las interacciones humanas con el entorno urbano, utilizando variables de actividades, entorno y ME. El resultado son segmentaciones o clústers que representan agrupaciones de características afines. Esto permite identificar patrones y subgrupos en el comportamiento humano, mejorando la comprensión de cómo las características del entorno influyen en las actividades recreativas. Los *insights* obtenidos ofrecen referencias valiosas para diseñar espacios que promuevan las actividades recreativas deseadas.

2.5 Modelado Predictivo

La variable de actividades se divide en cinco categorías y se transforma en variables dependientes binomiales. Se utiliza una red neuronal *feedforward* con un perceptrón multicapa, 32 capas ocultas y funciones de activación 'relu' y 'sigmoid' para predecir. El modelo se entrena con el 80% del *dataset* y prueba con el 20% restante, usando las ME como variables independientes y el tipo de actividad como variable dependiente.

El resultado es un modelo que predice actividades según métricas espaciales, útil tanto para evaluar diseños existentes como nuevas propuestas. Aunque tiene parámetros ajustables para mejorar la precisión, hay que evitar el sobreajuste para mantener su capacidad de predicción general. Cada modelo se entrena específicamente para cada actividad y se utiliza para las predicciones en el nuevo diseño a evaluar.

```
model <- keras_model_sequential() 
model %>%
layer_dense(units = 32, activation = 'relu', input_shape = ncol(x_train)) %>%
layer_dense(units = 1, activation = 'sigmoid') %>%
compile(loss = 'binary_crossentropy', optimizer = 'adam', metrics =
c('accuracy')) %>%
fit(x_train, y_train, epochs = 10, batch_size = 32, validation_split = 0.2)
```

2.6 Optimización y Diseño Generativo

Con el uso de Rhino y Grasshopper, se diseñó un canvas que contiene controladores paramétricos, generador de puntos aleatorios, voronoi, mass addition, merge y octopus para la optimización multiobjetivo.

La herramienta Octopus al igual que los algoritmos de optimización emplean la iteración para obtener el mejor resultado de una función objetivo o multiobjetivo como es el caso. En nuestro caso la optimización es a partir de modificar parámetros que logren dos objetivos: maximizar el área recreativa y la cantidad de edificios posibles. La solución óptima es un equilibrio ya que, a mayor cantidad de edificio menor es su área recreativa.

El área recreativa son áreas libres con vegetación y piso de Grass siendo el resultado de restarle al terreno las áreas de veredas, estacionamientos, alameda y la huella del edificio.

2.7 Evaluación del proyecto propuesto (tres escenarios y un caso referencial contrario)

La nueva propuesta es un prediseño optimizado con estructuras básicas para iniciar un diseño y su geometrización inicial. Se evaluará el prediseño para predecir y corroborar las actividades planificadas usando el indicador "interacciones positivas". Se extraen métricas espaciales con datos geoespaciales y se evalúan gráfica y cuantitativamente. El indicador mide la densidad de actividades recreativas predichas, calculada como la cantidad de observadores que predicen una actividad dividida por el total de observadores sintéticos.

$$\text{Ind} = \text{ObpPred} / \text{ObsTotal} \quad (1)$$

2.8 Validación del método MEPS

El indicador de la RSF es más alto que el del Conjunto Residencial Martinete sirviendo de referencia para los valores del indicador. La RSF muestra un diseño que favorece moderadamente las actividades recreativas al contrario con el Conjunto Residencial Martinete.

Se deben comprobar visualmente los tres escenarios para asegurar la coherencia con el diseño. Finalmente, se comparan los valores del indicador "interacciones positivas" entre los tres escenarios con los valores de referencia.

3 Resultados

A través de la revisión de datos de proyectos referenciales, se establecieron parámetros de diseño y se identificaron relaciones clave entre áreas recreativas y la densidad de edificios. Estos hallazgos proporcionan una perspectiva sobre cómo equilibrar la sostenibilidad con la viabilidad económica en el diseño urbano.

El Escenario 1 busca aumentar residentes y áreas libres, con un límite de 43 edificios; el Escenario 2 busca igualar las condiciones actuales con hasta

58 edificios; y el Escenario 3 explora la baja densidad con hasta 69 edificios, investigando la eficiencia del uso del terreno. Además, se utiliza el caso de Martinete Prueba, con parámetros similares al Escenario 3, como contraparte para una comparación adicional.

Tabla 2. Datos de los casos referenciales

Casos Referenciales	Área (Ha)	Densidad real (hab/Ha)	Residente actual	Densidad normativa	Máximo residentes
Residencial San Felipe	20.90	382.78	8,000	1,350	28,215
Residencial Martinete	1.89	899.47	1,700	900	1,701

Fuente: Elaboración propia, 2023

Tabla 3. Parámetros urbanos de Lima, aplicados a los tres escenarios y la propuesta de prueba que usa de base el terreno y parámetros del Conjunto Residencial Martinete

Propuestas	Densidad normativa (hab/Ha)	Máximo residentes	Nº Pisos edificio vivienda	Residentes por edificio de vivienda	Máximo número de edificios
RSF - Escenario 1 (RDA)	1,350	28,215	21	650	43
RSF - Escenario 2 (RDM)	900	18,810	11	325	58
RSF - Escenario 3 (RDB)	580	12,122	5	175	69
Martinete - Prueba MEPS (RDM)	580	12,122	11	175	69

Fuente: Elaboración propia, 2023

3.1 Del Análisis de Comportamiento, Entorno y Proximidad (ACEP)

Se identificaron seis clústers que reflejan características espaciales y tipos de actividades distintos. Estos clústers ofrecen una visión detallada de la interacción de los residentes con el entorno urbano, permitiendo comprender mejor las necesidades y comportamientos de los habitantes en diferentes contextos urbanos.

El Clúster 1 es un lugar polivalente con predominio en el paseo; el Clúster 2 es un sitio de descanso con equilibrio en su entorno; el Clúster 3 es una zona de tránsito; el Clúster 4 incluye actividades como pasear y jugar; el Clúster 5 es un patio de juegos con entorno mixto; y el Clúster 6 es similar al 2 pero con más conversaciones. Estos clústers (ver figura 1) muestran una variada combinación de intensidades de actividades y densidad del entorno, ofreciendo una comprensión profunda de cómo los residentes interactúan con el diseño urbano.

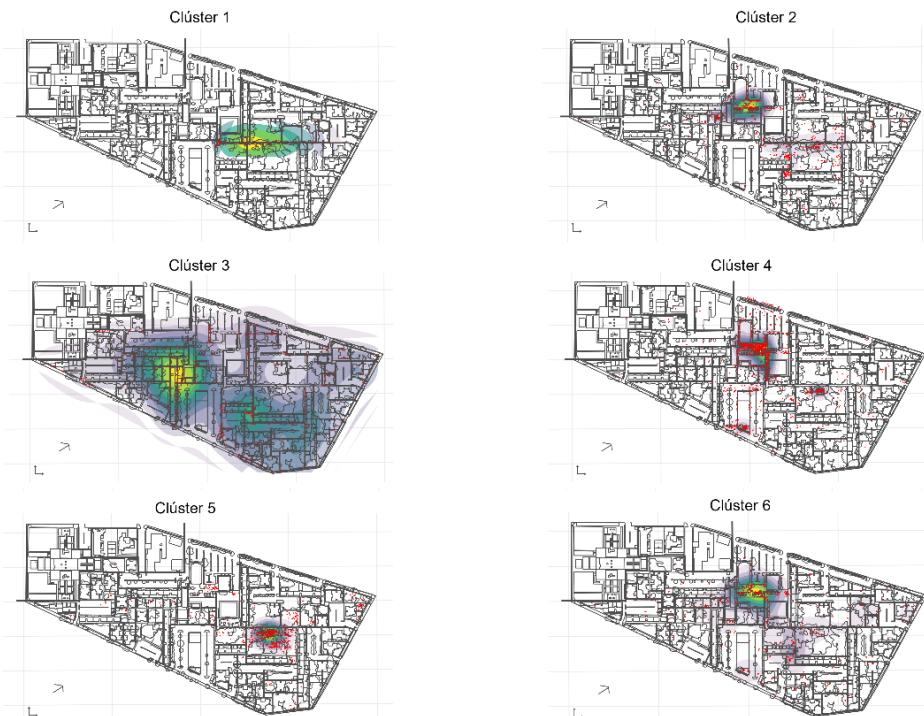


Figura 1. Mapa de calor de los seis agrupamientos por el comportamiento y características del entorno. Fuente: Elaboración propia, 2023

3.2 Del diseño generativo optimizado

El diseño generativo mostró que es posible aumentar significativamente el área recreativa sin comprometer la densidad habitacional. En particular, el método MEPS demostró ser eficaz al equilibrar estos dos aspectos, convergiendo hacia un diseño que maximiza el área recreativa mientras mantiene una densidad habitacional razonable.

En el caso de la RSF, las propuestas aumentaron el área recreativa del 33% al 60% y los residentes también aumentaron escalonadamente. Por ejemplo, en el Escenario 1, el máximo es de 43, y resolvió con 21 edificios para lograr un 68% de área recreativa.

En la Prueba del Conjunto Residencial Martinete, la optimización incrementó ligeramente la densidad habitacional con 11 torres, manteniendo el área recreativa en el rango del 60%. El algoritmo converger hacia el 60% del área recreativa, sin elevar demasiado la densidad ni reducir demasiado la cantidad de edificios.

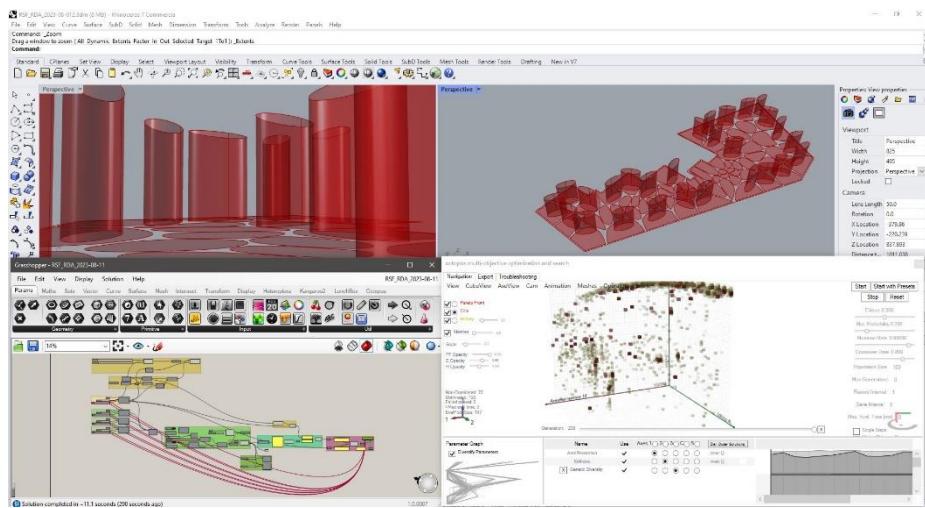


Figura 2. Resultado del diseño generativo, optimizado con la mayor área recreativa y la mayor cantidad de edificios. Fuente: Elaboración propia, 2023

Tabla 4. Resultados de la optimización con diseño generativo

Proyecto	Nº	Δ	Altura	Residentes	Hab/Ha	Grass	No Grass	% Grass
Residencial San Felipe	43	-	15	8,000	383	71,707	145,588	0.33
RSF - Escenario 1 (RDA)	21	22	21	13,650	628	148,367	68,928	0.68
RSF - Escenario 2 (RDM)	37	21	11	12,025	553	145,801	71,494	0.67
RSF - Escenario 3 (RDB)	57	12	5	9,975	459	131,426	85,869	0.60
Residencial Martínete	340	-	2	1,700	899	6,218	12,682	0.33
Martínete - Prueba MEPS (RDM)	11	58	11	1,925	1,019	11,762	7,138	0.62

Fuente: Elaboración propia, 2023

3.3 Del modelo predictivo

El análisis predictivo reveló que los diseños optimizados mejoran la predicción de actividades en comparación con el diseño original. Estos hallazgos sugieren que los diseños optimizados no solo son más sostenibles y económicamente viables, sino que también están más alineados con las necesidades y comportamientos de los residentes.

Los indicadores muestran un incremento en la predicción de las actividades entre la evaluación del diseño original versus los diseños optimizados (tres escenarios y la prueba MEPS).

Tabla 5. Resultados de los indicadores de la investigación

"interacciones positivas"	Paseo	Jugando	Charlando	Con niños	Descanso
Residencial San Felipe	0.08	0.12	0.08	0.10	0.04
RSF - Escenario 1 (RDA)	0.15	0.13	0.13	0.12	0.09
RSF - Escenario 2 (RDM)	0.11	0.15	0.08	0.12	0.04
RSF - Escenario 3 (RDB)	0.13	0.04	0.12	0.06	0.09
Residencial Martinete	0.08	0.06	0.02	0.01	0.01
Martinete - Prueba MEPS (RDM)	0.19	0.09	0.03	0.10	0.03

Fuente: Elaboración propia, 2023

En el caso de la RSF, el Escenario 1 muestra mejoras de casi al doble, el Escenario 2 tiene valores cercanos y el Escenario 3 es variado pero mejor que la referencia de control Residencial Martinete.

Tabla 6. Distribución de las actividades predichas

*Superficie de la actividad predicha (m ²)	Paseo	Jugando	Charlando	Con niños	Descanso
Residencial San Felipe	57,275	80,580	58,144	70,389	27,334
RSF - Escenario 1 (RDA)	105,228	91,561	88,954	85,478	59,329
RSF - Escenario 2 (RDM)	72,996	105,070	53,246	83,108	24,016
RSF - Escenario 3 (RDB)	90,376	30,731	83,977	41,159	60,830

*Cada observación positiva representa 79 m²

Fuente: Elaboración propia, 2023

De la distribución de áreas predichas el diseño del Escenario 1 aumenta al doble para los paseos, en el Escenario 2 desarrolla más para el juego y el Escenario 3 reduce más de la mitad para juegos. Para el caso de Martinete el diseño da énfasis a tener zonas para estar con niños.

Tabla 7. Distribución de las actividades predichas

**Superficie de la actividad predicha (m ²)	Paseo	Jugando	Charlando	Con niños	Descanso
Residencial Martinete	4,655	3,759	1,477	504	602
Martinete - Prueba MEPS (RDM)	11,361	5,495	2,065	6,006	1,596

**Cada observación positiva representa 7 m²

Fuente: Elaboración propia, 2023

Un patrón de diseño es colocar la zona de juegos al interior de las islas, bordeado por personas charlando y con niños; descansando, bordeando edificios y conexiones de tránsito verticales en la zona más densa.

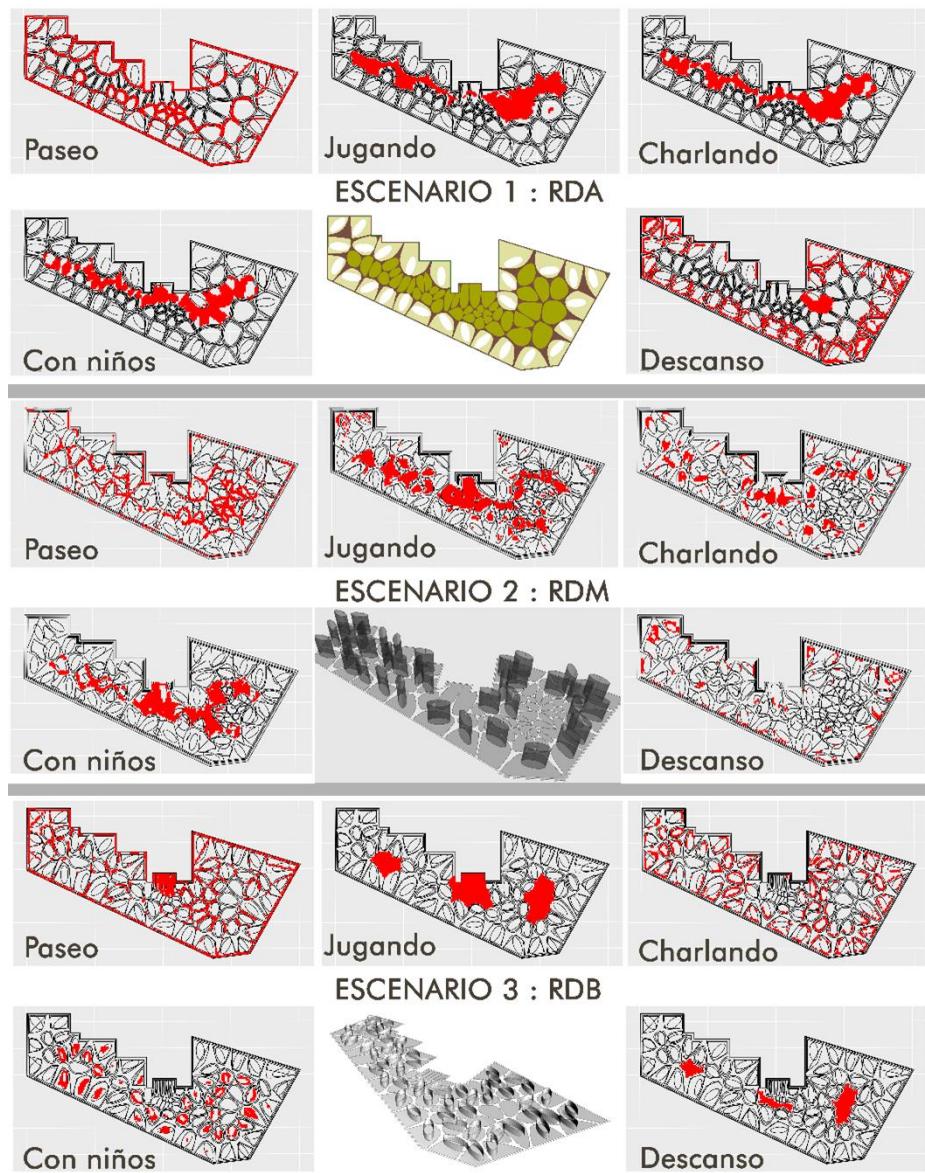


Figura 3. Mapeo de las actividades predichas de los tres escenarios. Fuente:
Elaboración propia, 2023

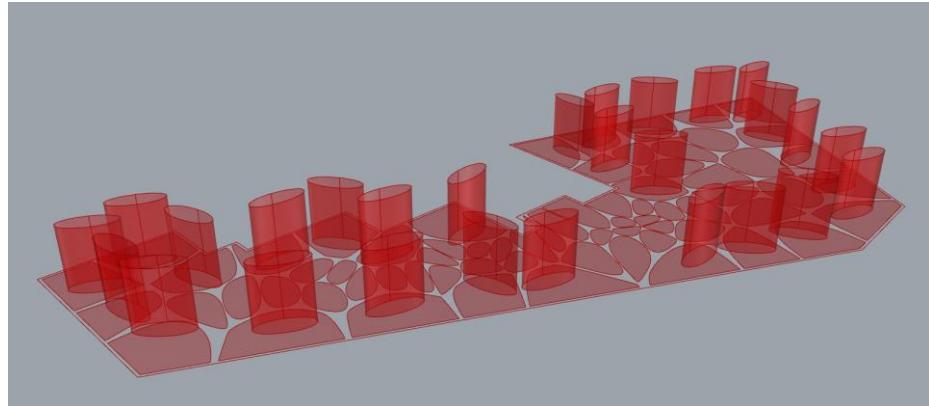


Figura 4. Vista del modelo básico, diseño paramétrico del Escenario 1. Fuente: Elaboración propia, 2023

Finaliza con una prueba del método MEPS con Martinete, explica correctamente las actividades predichas con lo real y en el caso propuesto es interesante la conectividad hacia el lado más agudo, de forma longitudinal.

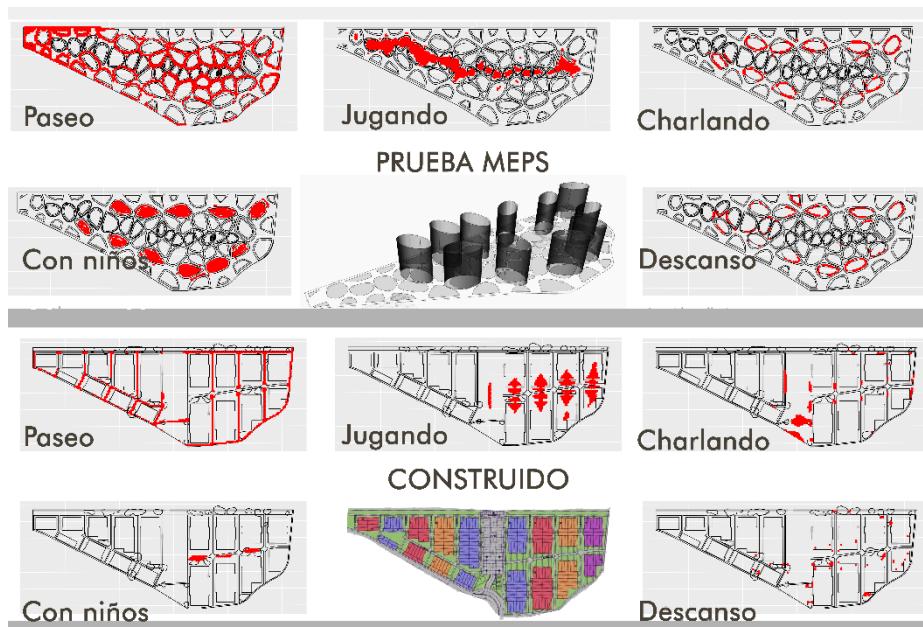


Figura 5. Mapeo de actividades del estado actual y de la propuesta de Prueba MEPS. Fuente: Elaboración propia, 2023

4 Discusión

Los resultados demuestran que es posible equilibrar la sostenibilidad con la viabilidad económica en el diseño urbano. El método MEPS, en particular, demostró ser una herramienta valiosa para lograr este equilibrio.

El análisis predictivo valida la eficacia de los diseños optimizados, reflejando las necesidades y comportamientos de los residentes.

Es esencial considerar estos hallazgos al desarrollar futuros proyectos urbanos, ya que ofrecen una hoja de ruta para crear espacios que no solo sean sostenibles y económicamente viables, sino que también satisfagan las necesidades y preferencias de los residentes.

Las predicciones no indican causalidad, sino correlaciones entre métricas, actividades y características del espacio, útiles para el diseño inicial.

Referencias

- Aghasi, N. (2018). Spatio-Temporal Analysis on Urban Traffic Accidents: A Case Study of Tehran City, Iran. *Journal of Geographic Information System*, 10(5), 532-555. <https://dx.doi.org/10.4236/JGIS.2018.105032>
- Grauwin, S., Sobolevsky, S., Moritz, S., Góder, I., & Ratti, C. (2014). Towards a comparative science of cities: using mobile traffic records in New York, London and Hong Kong. *arXiv preprint arXiv:1406.4400*. https://dx.doi.org/10.1007/978-3-319-11469-9_15
- Ma, S., Li, S., & Zhang, J. (2023). Spatial and deep learning analyses of urban recovery from the impacts of COVID-19. *Scientific Reports*, 11. <https://dx.doi.org/10.1038/s41598-023-29189-5>
- Mo, H., Zhang, Y., & Zhang, Y. (2022). Parametric Design and Spatial Optimization of East-West-Oriented Teaching Spaces in Shanghai. *Buildings*, 12(9), 1333. <https://dx.doi.org/10.3390/buildings12091333>
- Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1), 439-446. <https://doi.org/10.32614/RJ-2018-009>
- Wang, W., & Mo, X. (2014). Review on a Practical Approach of Sustainable Urban Design Strategy in the Perspective of Conflict in Shanghai. *International Review for Spatial Planning and Sustainable Development*, 2(4), 44-57. https://doi.org/10.14246/irspsd.2.4_44
- Yan, L., Chen, Y., Zheng, L., Zhang, Y., Liang, X., & Zhu, C. (2023). Intelligent Generation Method and Sustainable Application of Road Systems in Urban Green Spaces: Taking Jiangnan Gardens as an Example. *International Journal of Environmental Research and Public Health*, 20(4), 3158. <https://dx.doi.org/10.3390/ijerph20043158>
- Zhou, Y., Qian, C., Xiao, H., Xin, J., Wei, Z., & Feng, Q. (2019). Coupling Research on Land Use and Travel Behaviors Along the Tram Based on Accessibility Measurement—Taking Nanjing Chilin Tram Line 1 as an Example. *Sustainability*, 11(7), 2034. <https://dx.doi.org/10.3390/SU11072034>