



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FISICAS Y MATEMATICAS
DEPARTAMENTO DE INGENIERIA ELECTRICA**

**PREDICCIÓN DE DEMANDA DE PASAJEROS EN TIEMPO REAL PARA
EVALUACIÓN DE ESTRATEGIAS DE CONTROL EN SISTEMAS DE TRANSPORTE
PÚBLICO**

**TESIS PARA OPTAR AL GRADO DE MAGISTER EN CIENCIAS DE LA
INGENIERÍA, MENCIÓN ELECTRICA**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELECTRICISTA

FRANCISCO JAVIER VALENCIA QUILODRÁN

**PROFESOR GUÍA:
DORIS SÁEZ HUEICHAPAN**

**PROFESOR CO-GUÍA:
CRISTIAN CORTÉS CARRILLO.**

**MIEMBRO DE LA COMISIÓN:
GONZALO ACUÑA LEIVA**

**SANTIAGO DE CHILE
ENERO 2012**

RESUMEN DE LA TESIS PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELECTRICISTA Y AL GRADO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN ELÉCTRICA.
POR: FRANCISCO JAVIER VALENCIA QUILODRÁN
FECHA: 30/01/2012.
PROFESOR GUÍA: SRA. DORIS SÁEZ HUEICHAPAN

Predicción de demanda de pasajeros en tiempo real para evaluación de estrategias de control en sistemas de transporte público

El control en tiempo real es clave para reducir el atochamiento de los buses en servicios de transporte público, que se produce por la naturaleza estocástica de la demanda de pasajeros en paraderos. El control en sistemas de transporte público además permite regularizar el paso de los buses por paraderos, pero sobre todo permite disminuir los tiempos de espera de pasajeros.

En sistemas de transporte público, el control se ha desarrollado, generalmente, de forma heurística, lo que no permite reaccionar a cambios inesperados, como variaciones de la demanda, accidentes o aumento de congestión. Existen propuestas de control a través de estrategias de control predictivo híbrido (HPC) que tienen como ventaja valorar no solamente las acciones pasadas sino el comportamiento futuro del sistema. Estas estrategias consideran la demanda como una variable estocástica sobre la cual se aplican algoritmos de predicción. En este trabajo se presentan mejoras a la estrategia de control HPC mediante un carácter local del control para redes de gran tamaño y una nueva metodología de predicción de demanda basado en redes neuronales, lo cual permite disminuir los tiempos de espera y mejorar la regulación del paso de buses.

Se presenta una nueva metodología para utilizar los datos recolectados de un corredor real, facilitados por Transantiago S.A., y poder construir una demanda de pasajeros en paraderos confiable que represente en forma más precisa la realidad.

En relación a la predicción de demanda, se aplica un cambio de enfoque en la metodología, ya que no se busca predecir instantes de llegada de pasajeros, sino que cantidad de pasajeros que arriban en intervalos de tiempo. El disponer de un predictor basado en redes neuronales, permite mejores resultados, tanto en la medición de su error directo como en los indicadores de desempeño al utilizarlo en conjunto con el controlador.

Como trabajos futuros se propone implementar nuevas representaciones de sistemas de transporte público no diseñando redes circulares, sino que lineales con dos paraderos terminales y donde los paraderos en un sentido coinciden con paraderos en el sentido contrario. Esto sería de mucha utilidad en la comprensión de sistemas de transporte enmallados, donde por su tamaño, la aplicación de recorridos circulares no es lo más óptimo. Asociado a este punto se debe considerar la aplicación de estrategias de control donde el sistema de transporte tiene paraderos de transbordo que intercomunica dos recorridos.

Finalmente el bloque predictor necesario para la aplicación del control HPC, se puede mejorar buscando algoritmos de predicción que consideren el historial no solamente del paradero consultado, sino que también utilicen la información de sus vecinos.

Índice

1. Introducción	7
1.1. Motivación	7
1.2. Objetivos	8
1.3. Estructura de la tesis	8
2. Análisis de demanda en un sistema de transporte público real	10
2.1. Descripción y obtención de la demanda	10
2.1.1. Información de pasajeros en paraderos obtenidos de datos de Transantiago S.A.	10
2.1.2. Estimación instantes de llegada de pasajeros basado en datos de Transantiago S.A.	20
2.1.3. Estimación destinos de pasajeros simulados	21
2.2. Análisis de demanda simulada	21
2.3. Generación de nuevos datos simulados (réplicas)	26
2.4. Discusión	27
3. Estrategia de control en tiempo real para sistemas de transporte público	28
3.1. Introducción	28
3.2. Formulación del problema	32
3.3. Estrategia de control	35
3.4. Restricciones operacionales	40
3.5. Función objetivo	41
3.6. Modelo de predicción de la demanda	43
3.7. Método de solución del controlador predictivo híbrido basado en estrategia de todos los paraderos	44
3.8. Método de solución del controlador predictivo híbrido basado en estrategia de selección de paraderos	45
3.9. Análisis de sensibilidad de métodos de solución	46

3.9.1.	Análisis de tamaño de selección de paraderos P^*	48
3.9.2.	Análisis de pasos futuros equivalentes	50
3.9.3.	Análisis de <i>headway</i> de referencia	52
3.9.4.	Análisis de <i>headway</i> de referencia según la flota óptima	53
3.9.5.	Análisis de variación de pesos θ en función objetivo	54
3.10.	Discusión	63
4.	Análisis de robustez de la estrategia de control en tiempo real con respecto a la demanda	65
4.1.	Descripción de escenarios	65
4.2.	Resultados por simulación	68
4.3.	Discusión	71
5.	Modelos de predicción de demanda origen-destino (OD)	72
5.1.	Estimación de demanda en sistemas de transporte público	72
5.2.	Estimación y predicción de demanda en sistema de gestión de tráfico	73
5.3.	Propuesta de predicción de cantidad de pasajeros en paradero	76
5.3.1.	Modelo heurístico basado en promedios para tasa de llegada constante	77
5.3.2.	Modelo de predicción de cantidad de pasajeros con redes neuronales (NN)	78
5.3.3.	Error de predicción directo	81
5.4.	Análisis y sensibilización de parámetros para modelo de predicción heurístico	84
5.5.	Aplicación de predictores a red de transporte público	94
5.6.	Discusión	96
6.	Conclusiones y trabajos futuros	97
	Referencias	118

Índice de figuras

1.	Gráfico cartesiano del recorrido 506 ida	12
2.	Proyección de paradero a ruta de referencia	14
3.	Posición buses entre 7:00 y 10:00	15
4.	Distribución de transacciones en ruta	16
5.	Curva de carga de paradero P=4 a intervalo variable	18
6.	Curva de carga de paradero P=53 a intervalo variable	18
7.	Curva de carga de paradero P=21 a intervalo variable	19
8.	Volumen de paraderos	23
9.	Curvas de carga agregada	24
10.	Matriz origen-destino día 2	26
11.	Modelo de corredor circular	32
12.	Diagrama de estrategia control predictivo	34
13.	Ejemplo de velocidad de bus entre paradas consecutivas	37
14.	Variación tiempo de espera (a), Regulación promedio (b), Tiempo <i>holding</i> total (c), Tiempo de <i>skipping</i> total (d) y Valor de la función objetivo (e).	49
15.	Variación tiempo de espera (a), Regulación promedio (b), Tiempo <i>holding</i> total (c), Tiempo de <i>skipping</i> total (d) y Valor de la función objetivo (e).	51
16.	Variación tiempo de espera (a) y valor función objetivo (b)	53
17.	Variación tiempo de espera (a) y valor de función objetivo (c) para flota óptima 10; y Variación tiempo de espera (b) y valor de función objetivo (d) para flota óptima 12.	54
18.	Variación tiempo de espera (a), Regulación promedio(b), Tiempo <i>holding</i> total(c), Tiempo de <i>skipping</i> total(d) y Valor de la función objetivo (e).	56
19.	Variación tiempo de espera (a), Regulación promedio(b), Tiempo <i>holding</i> total(c), Tiempo de <i>skipping</i> total(d) y Valor de la función objetivo (e).	58
20.	Variación tiempo de espera (a), Regulación promedio(b), Tiempo <i>holding</i> total(c), Tiempo de <i>skipping</i> total(d) y Valor de la función objetivo (e).	60

21.	Variación tiempo de espera (a), Regulación promedio(b), Tiempo <i>holding</i> total(c), Tiempo de <i>skipping</i> total(d) y Valor de la función objetivo (e). . . .	62
22.	Comparación Matrices OD según la forma de la carga	67
23.	Comparación Matrices OD según volumen de la carga	67
24.	Coefficiente de Lipschitz para diferentes valores de regresor	80
25.	$Error_p$ agregado según predictor	83
26.	$Error_p$ para distintos paraderos	84
27.	Tiempo de espera total pasajeros (a); Regulación promedio (b); Tiempo de holding total (c); Tiempo de skipping total (d) y Valor de la función objetivo (e).	86
28.	Tiempo de espera total pasajeros (a); Regulación promedio (b); Tiempo de holding total (c); Tiempo de skipping total (d) y Valor de la función objetivo (e).	89
29.	Tiempo de espera total pasajeros (a); Regulación promedio (b); Tiempo de holding total (c); Tiempo de skipping total (d) y Valor de la función objetivo (e).	91
30.	Tiempo de espera total pasajeros (a); Regulación promedio (b); Tiempo de holding total (c); Tiempo de skipping total (d) y Valor de la función objetivo (e).	93
31.	Variación tiempo de espera pasajeros (a); Regulación promedio (b); Tiempo de holding total (c); Tiempo de skipping total (d) y Valor de la función objetivo (e).	95
32.	Procedimiento básico de identificación de estructura	106
33.	Una entrada es aplicada a el sistema y se observa cómo la salida es afectada	107
34.	Modelo sistema dinámico	109
35.	Red Neuronal	111

Índice de tablas

1.	Transacciones en buses	11
2.	Recorrido de buses	11
3.	Posiciones de paraderos	11
4.	intervalos de tiempo y cantidad de pasajeros	17
5.	Instantes de llegada para un paradero particular	21
6.	diferencia entre días disponibles para paraderos	25
7.	<i>Headways</i> de referencia analizados	52
8.	Pesos de análisis de tiempo de espera	55
9.	Pesos de análisis de componente <i>holding</i>	59
10.	Pesos de análisis de componente <i>skipping</i>	61
11.	Descripción escenarios	66
12.	Índices de desempeño para escenarios 1 y 2	69
13.	Índices de desempeño para escenarios 3 y 4	69
14.	Índices de desempeño para escenarios 5 y 6	69
15.	Índices de desempeño para escenarios 7 y 8	70
16.	Valores de parámetros	85
17.	Valores de parámetros	88
18.	Valores de parámetros	90
19.	Valor componente α	90
20.	Valores de parámetros	92
21.	Tamaño d_{1a}	92
22.	Formato de los datos online	101
23.	Datos de un día Ω cualquiera	103

1. Introducción

1.1. Motivación

Cuando no se tiene un sistema de control en una red de transporte público, tiende a ocurrir atochamiento de buses por la naturaleza estocástica de la demanda de pasajeros en los paraderos. Además se incrementa la variabilidad de la frecuencia de paso de los buses, pero la peor consecuencia es el empeoramiento de los tiempos de espera de los pasajeros. Esta situación daña la percepción del atochamiento de servicio entregado por parte de los clientes, Ortúzar (1983).

En las redes de transporte público los sistemas de control se han desarrollado, en general, de forma heurística, lo que no les permite reaccionar en tiempo real a cambios inesperados como variaciones de la demanda, accidentes o aumento de congestión. Esto genera que la frecuencia de los buses sea muy variable lo que afecta negativamente los tiempos de espera y viaje de los pasajeros.

Existen propuestas de control en tiempo real para redes de transporte público a través de estrategias de control predictivo híbrido (HPC) que tienen como ventaja valorar no solamente las acciones pasadas sino el comportamiento futuro del sistema. Estas estrategias consideran la demanda como una variable estocástica sobre la cual se aplican algoritmos de predicción simples.

En este trabajo se busca mejorar el control de redes de transporte público a través de la aplicación de estrategias de control HPC de carácter local para redes de gran tamaño, y por otro lado, analizar el impacto de la demanda en los algoritmos de control, considerando la precisión de los algoritmos de predicción y observando su influencia en los modelos de control. El análisis de demanda en sistemas de transporte ha estado enfocado a dos áreas: el análisis de demanda en forma estática a través de las matrices de origen destino para planificación de redes de transporte y el análisis de demanda en redes de transporte en general para el estudio del flujo de vehículos para la gestión de tráfico, por lo que el análisis de demanda de pasajeros en tiempo real en sistemas de transporte público es un tema no tratado en la

literatura. Para desarrollar los estudios de control de redes, se propone trabajar sobre los datos facilitados por Transantiago S.A. y obtener información sobre el comportamiento de la demanda real con la cual generar un simulador más cercano a la realidad.

Para predecir la demanda, esta propuesta se basa en un modelo de redes neuronales, que es la herramienta habitual para predicción de sistemas no lineales.

1.2. Objetivos

El objetivo general de esta tesis es evaluar el impacto de la demanda de pasajeros en la aplicación de estrategias de control de sistemas de transporte público, considerando una nueva propuesta de método de solución del controlador y un nuevo algoritmo de predicción. El desarrollo de esta evaluación se realiza sobre un simulador que modela una demanda de pasajeros basada en datos reales de un sistema de transporte público.

Por lo tanto, los objetivos específicos de este trabajo son:

- Proponer un nuevo método de solución local para el algoritmo de control predictivo para redes de transporte de gran tamaño.
- Modelar una demanda de pasajeros adaptada a la realidad basada en datos de sistemas de transporte público reales.
- Diseñar un método de predicción de instantes de tiempo de llegada de pasajeros para sistemas de transporte público muestreados por eventos.
- Evaluar el comportamiento de los algoritmos de control al someterlos a cambios en la demanda.

1.3. Estructura de la tesis

Este trabajo se encuentra estructurado del siguiente modo. En el capítulo 2 se presenta un análisis sobre una nueva fuente para extraer un modelo de demanda de pasajeros que sea capaz de reflejar de forma más precisa el comportamiento real de un sistema de transporte público.

A continuación, en el capítulo 3, se analiza una nueva estrategia de control HPC basado en selección de paraderos que permite obtener indicadores de desempeño adecuados cuando se trabaja en sistemas de transporte con alta cantidad de paraderos y con una distancia de circuito considerable.

Para observar la robustez de esta nueva estrategia frente a diferentes tipos de demanda de pasajeros, en el capítulo 4 se analiza el rendimiento de la estrategia de control aplicado a un sistema de transporte variando las características de la demanda.

En el capítulo 5 se presenta el análisis sobre metodologías de predicción de demanda que se utiliza en la aplicación de estrategias de control HPC. Este punto adquiere relevancia al disponer de una demanda de pasajeros que refleja de mejor forma el comportamiento real de un sistema de transporte público..

Finalmente en el capítulo 6 se presentan las conclusiones finales y los trabajos futuros.

2. Análisis de demanda en un sistema de transporte público real

Para implementar estrategias de control para transporte público se requiere disponer de un sistema para realizar las simulaciones donde un elemento decisivo es la existencia de una demanda que refleje lo más fielmente posible una situación real. Para lograr esta característica se utilizan los datos facilitados por Transantiago S.A. sobre el funcionamiento de la red de transporte de la ciudad de Santiago. A continuación se presenta el procedimiento para trabajar sobre estos datos.

2.1. Descripción y obtención de la demanda

La demanda del sistema corresponde a la llegada de los usuarios a los distintos paraderos a lo largo de la ruta. Esta llegada varía según el paradero y según la franja horaria en que el sistema se encuentre operando. Con la llegada de cada usuario se tiene como dato su hora de arribo y su paradero de destino.

Los instantes de llegada se basan en datos del Transantiago S.A. y los destinos se simulan según una distribución Gaussiana.

A continuación se presentan los pasos necesarios para la obtención de la demanda.

2.1.1. Información de pasajeros en paraderos obtenidos de datos de Transantiago S.A.

Los instantes de llegada se obtienen a través de un estudio de la demanda facilitada por Transantiago S.A. que consiste en los datos generados por el uso de la tarjeta de pago del sistema de transporte público de Santiago. Esta información está disponible para las estaciones de pago¹ y en los buses mismos, donde los pasajeros cancelan al subir (tabla 1). Además se entrega la ubicación de los buses cada cierto tiempo dentro del recorrido que realizan (tabla 2). Finalmente se tiene la posición geográfica de todos los paraderos y un

¹Puntos para cancelar disponibles en ciertos paraderos de alta demanda.

recorrido referencial (tabla 3).

tipo	bus	tiempo
BUS	BFKB-70	2008-09-03 06:43:12
BUS	BFKB-70	2008-09-03 06:43:16
BUS	BFKB-70	2008-09-03 06:45:28
BUS	BFKB-70	2008-09-03 06:45:29
BUS	BFKB-70	2008-09-03 06:45:30
BUS	BFKB-70	2008-09-03 06:45:38
BUS	BFKB-70	2008-09-03 06:45:39
BUS	BFKB-70	2008-09-03 06:45:56
BUS	BFKB-70	2008-09-03 06:49:06
BUS	BFKB-70	2008-09-03 06:49:07
BUS	BFKB-70	2008-09-03 06:49:10

Tabla 1: Transacciones en buses

bus	tiempo	x	y
BFKB-70	2008-09-01 04:46:26	-9197.68766638002	-6844.62656628527
BFKB-70	2008-09-01 04:46:56	-8980.50452969101	-6657.65907219518
BFKB-70	2008-09-01 04:47:26	-8828.93015558401	-6528.95092811529
BFKB-70	2008-09-01 04:47:56	-8599.77518506901	-6336.70914248563
BFKB-70	2008-09-01 04:48:26	-8228.940105549	-6015.94543057587
BFKB-70	2008-09-01 04:48:56	-8045.68666029698	-5832.12184893526
BFKB-70	2008-09-01 04:49:26	-8099.381648226	-5717.94506940525
BFKB-70	2008-09-01 04:49:56	-8024.48516218597	-5662.92293095589
BFKB-70	2008-09-01 04:50:26	-7846.02872720297	-5529.54046092555
BFKB-70	2008-09-01 04:50:56	-7638.18970637198	-5373.2330438653
BFKB-70	2008-09-01 04:51:26	-7430.34722409502	-5299.86707108561
BFKB-70	2008-09-01 04:51:56	-7303.40769972099	-5428.45510516595
BFKB-70	2008-09-01 04:52:26	-7298.991900853	-5188.45625714585
BFKB-70	2008-09-01 04:52:56	-7298.99291711696	-4938.80042272527
BFKB-70	2008-09-01 04:53:26	-6953.48656399397	-4957.03902509529
BFKB-70	2008-09-01 04:53:56	-6936.62980046502	-4781.1474661557
BFKB-70	2008-09-01 04:54:26	-6943.53779447899	-4499.87390226591
BFKB-70	2008-09-01 04:54:56	-6961.38946140802	-4325.41879944596
BFKB-70	2008-09-01 04:55:26	-6980.05419251201	-4103.16227914579

Tabla 2: Recorrido de buses

id	x	y
1	-9146.9	-6758.1
2	-8920.9	-6581.8
3	-8771.7	-6459.2
4	-8578.2	-6301.3
5	-8366	-6113.3
6	-8129.6	-5896.3
7	-8093.9	-5683.1
8	-7857.7	-5499.1
9	-7622.2	-5322.2
10	-7423.5	-5358.4

Tabla 3: Posiciones de paraderos

Para este trabajo se seleccionan los datos de la línea 506 solamente de ida ya que la propuesta es sobre un sólo corredor y no una red de líneas.

La línea 506 ida contiene un total de 85 paraderos en un recorrido que comienza en la comuna de Maipú y finaliza en la comuna de Peñalolén, con un largo total de 33,2 kilómetros. En la Figura 1 se aprecia gráficamente el recorrido con sus paraderos ajustado para evitar tener la posición en UTM (Universal Transverse Mercator), por eso aparecen posiciones negativas.

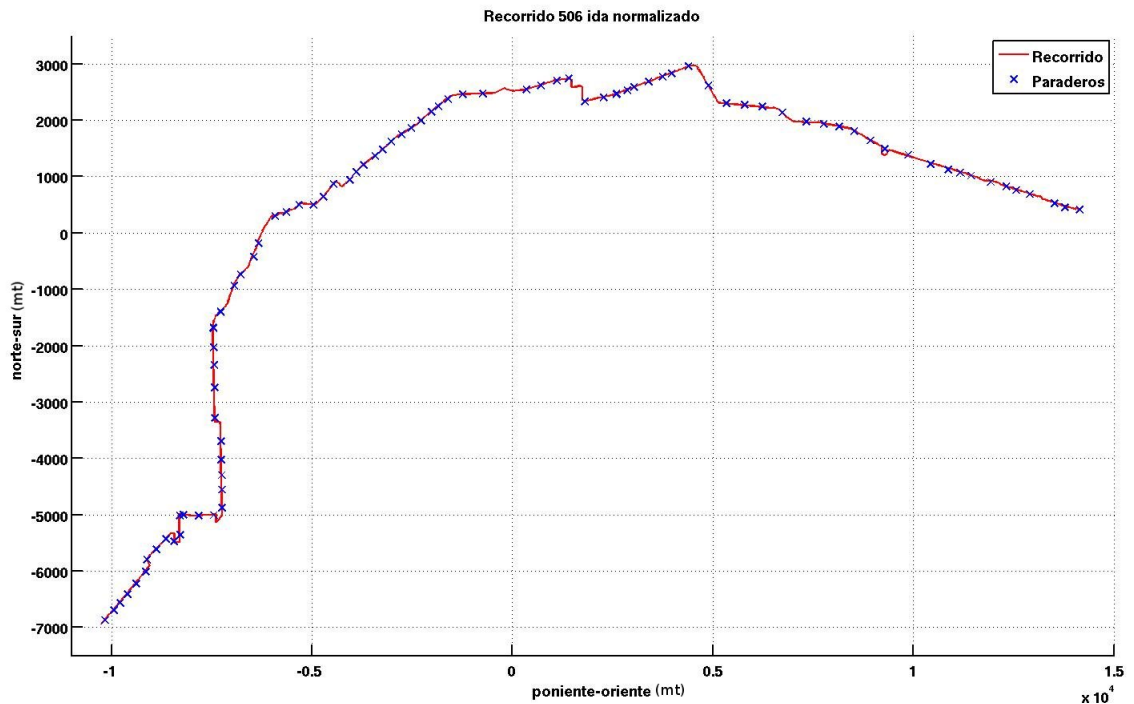


Figura 1: Gráfico cartesiano del recorrido 506 ida

Los datos facilitados corresponden a la semana del 1 al 7 de septiembre de 2008. En esta semana se realizaron 185,752 transacciones bips en buses y 37,674 transacciones bips en zonas pagas. Los datos de los buses corresponden a 73 buses diferentes con 364,213 posiciones. Todos los datos disponibles se normalizan para disminuir la influencia de las dimensiones en los cálculos. Además se filtran los posibles puntos que no pertenezcan al dominio de estudio. Los datos disponibles por las transacciones en los buses no permiten obtener el instante de llegada de los pasajeros a los paraderos, ya que solamente se obtiene información de un pasajero cuando marca una transacción, lo cual generalmente es cuando se sube al bus y no

cuando llega al paradero, que sería lo ideal. Considerando esta limitante, lo que si se puede obtener es la cantidad de personas que llegan a un paradero dentro de intervalos definidos por la llegada de los buses para después estimar los instantes de llegada.

Complementando lo anterior, en las zonas pagas, la información que se dispone son los instantes de llegada pero no se tiene cuál es el bus que abordan, por lo que se estima una frecuencia de paso de todas las líneas por paradero para distribuir las transacciones en forma proporcional a cada línea.

A continuación se presenta el procedimiento aplicado a los datos de transacciones en buses para obtener las cotas superior e inferior de los intervalos y la cantidad de pasajero que llegan dentro de éste.

Procedimiento para obtención información de paraderos

1. **Cambiar variable de ruta de referencia:** Para no tener que trabajar con dos dimensiones en la posición, se define utilizar la distancia hacia el origen de la ruta como variable que indica la posición. Esta distancia se debe calcular a través de la ruta referencial del recorrido y no como la distancia entre dos puntos en un plano cartesiano, lo que permite mantener la consistencia de los datos. Como ruta de referencia se entiende un conjunto de puntos fijos que representan el recorrido ideal de los buses.
2. **Proyectar posiciones a ruta referencial:** Se proyectan todas las posiciones en el sistema (posición GPS de buses, ubicación de los paraderos) a la ruta referencial. Esta proyección se realiza considerando la mínima distancia del punto a proyectar al segmento más cercano de la ruta referencial. En la Figura 2 se aprecia la proyección de un paradero a la ruta referencial.

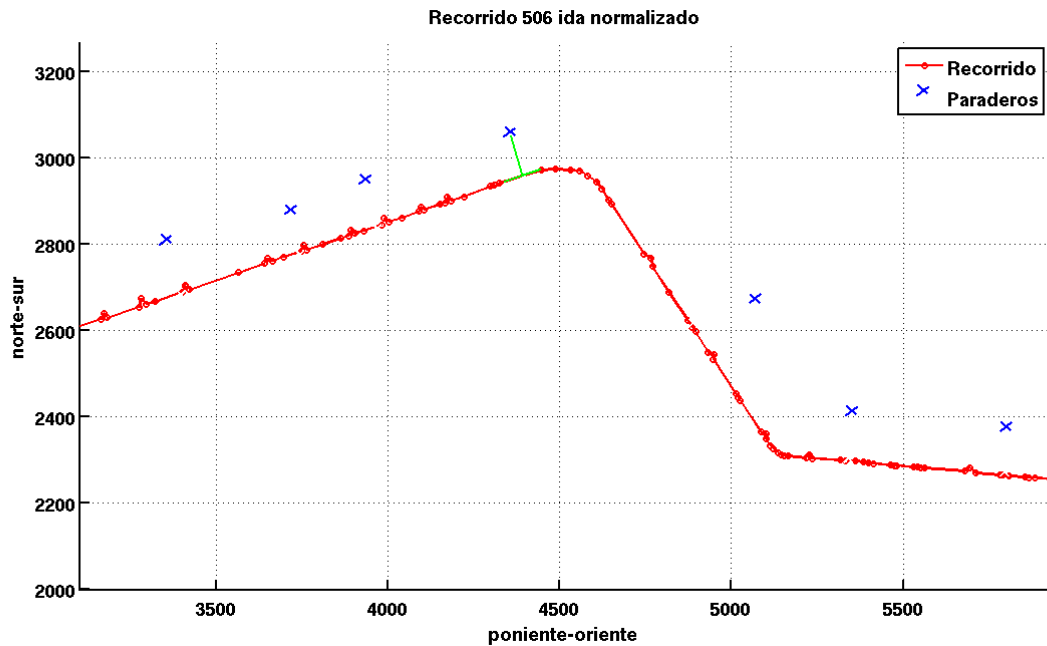


Figura 2: Proyección de paradero a ruta de referencia

En este paso se elimina un paradero porque su posición no calza con la ruta referencial. Esto se justifica porque la ruta referencial no considera trabajos realizados a posteriori que implican una modificación de la ubicación de paraderos y rutas de buses.

En la proyección de las posiciones GPS se eliminan 70,414 datos, lo que equivale a un 19.33% de datos. A pesar de ser un valor considerable se debe entender que es solamente la asociación de un bus a un punto en el recorrido y no la asociación de las transacciones a los paraderos que es el objetivo final.

3. **Asignar tiempo de paso de bus por paraderos:** Se estima el tiempo de paso de los buses por los paraderos. Si dos posiciones consecutivas de GPS de un bus tienen entre ellas la posición de un paradero, se proyecta el tiempo paso del bus por ese punto en forma proporcional según la distancia (se asume una velocidad constante promedio entre puntos). Se ordenan por paradero para poder visualizar el paso de los buses y determinar los intervalos de tiempo entre los cuales llegan los pasajeros. En la Figura 3 se presenta el resultado para el día 1 entre 7:00 y 10:00.

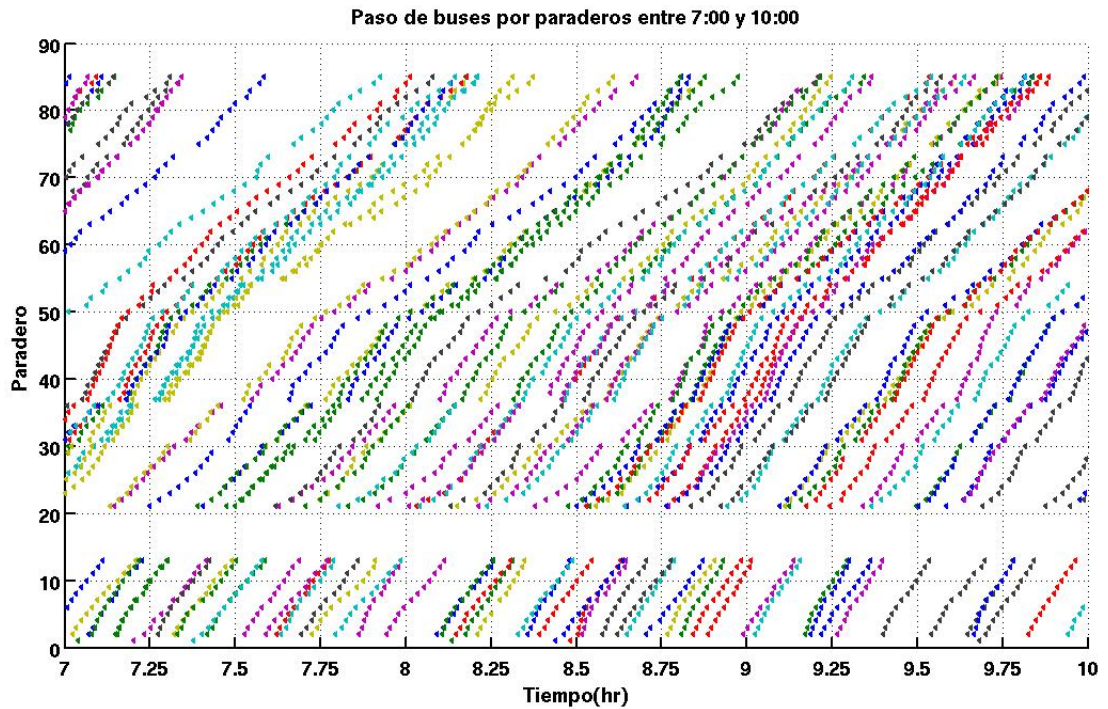


Figura 3: Posición buses entre 7:00 y 10:00

Las zonas vacías son paraderos que se cambiaron después de definir el recorrido de referencia, por lo que los datos de GPS aparecen alejados de la ruta de referencia y por ende no se pueden asociar a los paraderos disponibles. Se aprecian claramente las zonas de congestión, donde se observan los buses avanzando sin ninguna espera entre ellos, además de como para algunos paraderos pueden pasar 3 buses seguidos para después tener intervalos de tiempo de más de 20 minutos.

4. **Asociar transacciones bips a paraderos:** Las transacciones bips se deben asociar a un paradero donde se estima más probablemente que el pasajero se subió al bus. Esta asociación se realiza asignando primero a cada transacción una distancia al paradero terminal, considerando las dos posiciones del bus donde se subió que están antes y después de la transacción, para después proyectar proporcionalmente según el tiempo la distancia representativa.

Este cambio a distancia desde el origen refleja claramente la pertenencia de las transacciones a paraderos específicos, como se aprecia en la Figura 4. Además se ob-

serva que una parte de los paraderos no tienen asociados transacciones lo que valida no considerar todos los paraderos para el simulador, sino solamente los que tienen una demanda significativa.

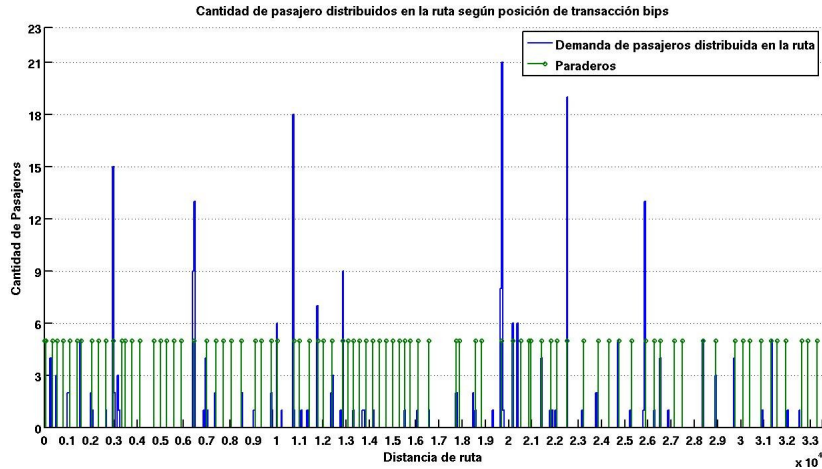


Figura 4: Distribución de transacciones en ruta

Este paso genera una pérdida de 24,898 transacciones que no se pueden asignar lo que es un 13,4 % de los datos totales disponibles.

A continuación se define un rango de 50 mts. alrededor de cada paradero, y todas las transacciones que estén dentro de este rango se asocian al paradero. Este paso tiene una pérdida de 29,056 transacciones que es un 15.64 % de los datos totales disponibles, estos datos se descartan por estar muy alejados de los paraderos ya que pueden ser pasajeros que se suben a la mitad de un recorrido, situación que no se espera que suceda o se deba considerar.

Finalmente de las 185,752 transacciones quedan 131,793 para poder ser asignadas a los intervalos de tiempo.

5. **Asignar transacciones a intervalos de tiempo:** Que un pasajero haya marcado en un bus i , significa que debe haber arribado antes de la llegada estimada de ese bus al paradero p , pero después de que el bus $i - 1$ pasara por paradero p (Se asume que los pasajeros logran subirse al primer bus). Entonces ese pasajero debe haber llegado en algún momento entre los instantes de tiempos definidos por el paso de los dos buses

anteriores $t_{i,p}$ y $t_{i-1,p}$, así que se asigna a ese intervalo de tiempo. Esta asignación nuevamente deja datos afuera, descartando 22,574 transacciones que es equivalente a un 12.15 %.

Después de realizar los pasos anteriores se obtiene la información más real sin recurrir a simulaciones logrando trabajar sobre 109,224 transacciones que equivale a un 58.8 % (ejemplo en tabla 4). A pesar de que el porcentaje de descarte es alto, esto ocurre por arrastre de errores iniciales que no permiten finalmente asignar transacciones a paraderos, por lo que este error se encuentra condensado a paraderos específicos, lo que motiva el hecho de no trabajar con 86 paraderos para el simulador, sino que realizar una selección donde realmente hayan datos significativos.

Inicio intervalo (hr)	Fin intervalo (hr)	Cantidad pasajeros
6:55	6:57	25
6:57	7:02	286
7:02	7:08	393
7:08	7:11	139
7:11	7:12	10
7:12	7:18	84

Tabla 4: intervalos de tiempo y cantidad de pasajeros

En las Figuras 5, 6 y 7 se presentan las curvas de carga para un día lunes de una selección de paraderos que refleja la buena obtención de información realizada anteriormente. Esta es la información más directa que se puede obtener.

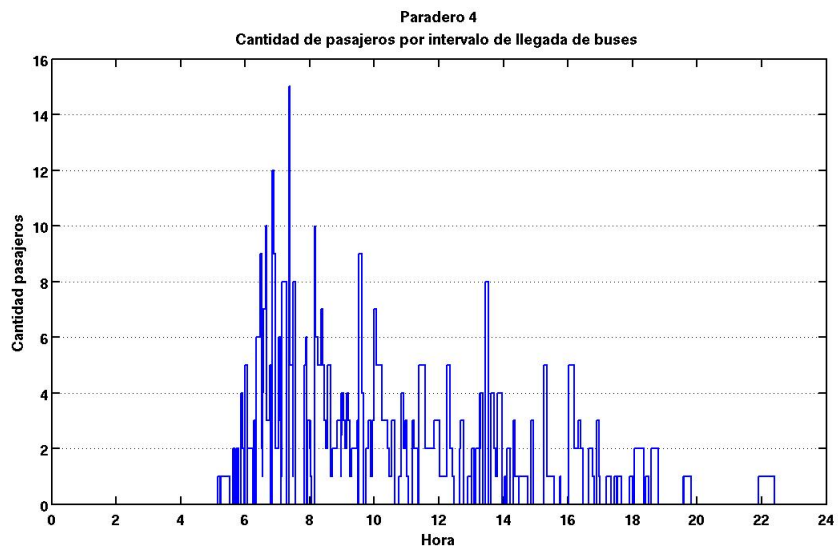


Figura 5: Curva de carga de paradero P=4 a intervalo variable

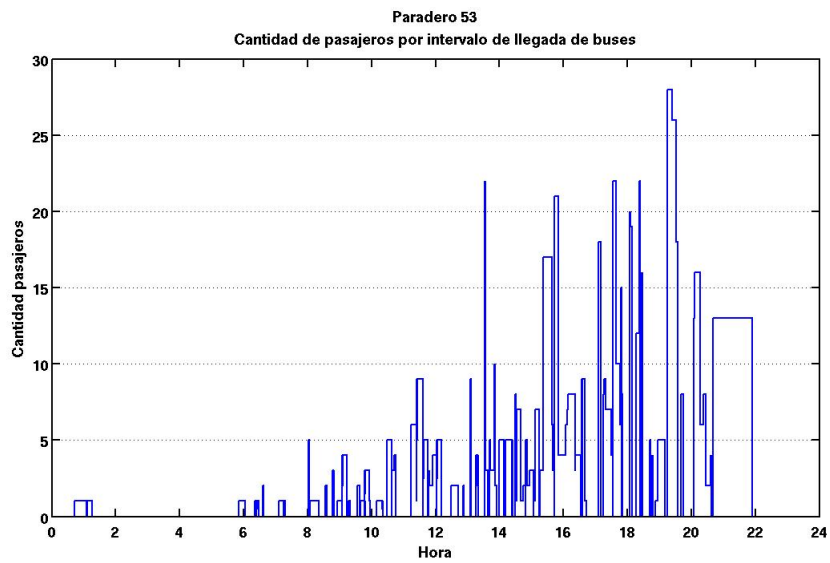


Figura 6: Curva de carga de paradero P=53 a intervalo variable

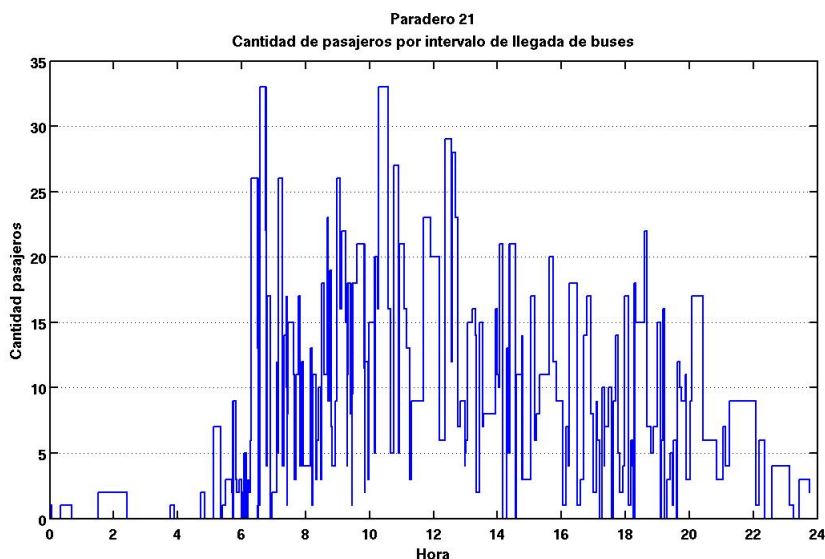


Figura 7: Curva de carga de paradero P=21 a intervalo variable

Se observa claramente diferentes tipos de paraderos, dependiendo de cuando se producen los peak de demanda. Se tienen paraderos tipo Figura 5 que representa el caso cuando la demanda máxima ocurre en el periodo de la mañana para después decaer. En el caso de la Figura 6 el peak ocurre en la tarde; finalmente en la Figura 7 se tiene una demanda más o menos homogénea a través del día.

Estos datos no se pueden usar directamente como entrada en los simuladores, ya que se necesita el tiempo de llegada de los pasajeros al paradero y no su valor agregado dentro de un intervalo variable. En la sección 2.1.2 se muestran los pasos aplicados para obtener los instantes de llegada. A continuación se presenta la otra fuente de transacciones, que son las realizadas en las zonas pagas.

Obtención información zonas pagas. En el sistema existen ciertos paraderos donde en vez de cancelar el pasaje cuando se suben al bus, se cancela en el instante en que se llega al paradero, estos paraderos se conocen como *Zonas pagas*. Este sistema de paraderos se implementa en puntos de la red con alta demanda, donde el tiempo de boarding es muy alto por una sola puerta lo que enlentece el desplazamiento de los buses, además estas aglomeraciones incitan a la evasión del pago del pasaje. La implementación consiste en cercar el alrededor del paradero y colocar en las puertas de acceso del paradero validadores. Como

consecuencia se permite la subida de pasajero por todas las puertas, lo que acelera el recorrido del trayecto.

En el caso particular del recorrido 506 ida se tienen 3 paraderos que funcionan como zonas pagas, eso si solamente en ciertos horarios (en las puntas mañana o en las puntas tarde) de 06:30 a 10:00 y de 17:00 a 21:00.

Los datos disponibles en este caso corresponden exactamente a los instantes de llegada de los pasajeros, que es la información que se está buscando estimar, pero se tiene como problema que no se sabe a cuál de los buses que pasan por ese paradero se sube cada pasajero. Para poder asignar los buses a los pasajeros (y de esa forma estimar la cantidad de pasajeros que corresponden al recorrido analizado), se prorroga la cantidad total de pasajeros según la frecuencia de paso de los diferentes recorridos en cada paradero y se asume una llegada homogénea según buses de los pasajeros. En este caso se tienen otros 5 recorridos que pasan en alguno de los 3 paraderos.

La información de las zonas pagas se utiliza para enriquecer los datos obtenidos en el procedimiento de los paraderos comunes.

Considerando ambas fuentes (zonas pagas y paraderos comunes) se obtienen los instantes de llegada de pasajeros por paradero.

2.1.2. Estimación instantes de llegada de pasajeros basado en datos de Transantiago S.A.

En la sección 2.1.1 se presenta cómo obtener información sobre la demanda de pasajeros desde los datos entregados por Transantiago S.A. Esta información es de dos tipos, en los paraderos que no tienen zonas pagas, se dispone de intervalos de tiempo variables y la cantidad de pasajeros que llegan dentro de cada uno de esos intervalos, y en los paraderos con zonas pagas se dispone de instantes de llegada de pasajeros al paradero.

Para obtener instantes de llegada de pasajeros en paraderos sin zonas pagas, se asume una llegada homogénea de pasajeros dentro de cada intervalo disponible. Esto se une a lo disponible en los paraderos con zonas pagas, disponiendo finalmente de listados de instantes

de tiempo que representan la llegada de un pasajero al paradero respectivo, como se aprecia en la tabla 5.

Instante de llegada (hr)
07:10:09
07:10:19
07:10:29
07:10:39
07:10:49
07:10:59
07:11:09
07: 11:20

Tabla 5: Instantes de llegada para un paradero particular

2.1.3. Estimación destinos de pasajeros simulados

Como se ha indicado previamente, la información facilitada por Transantiago S.A. no permite definir directamente el destino de los pasajeros. Entonces para poder disponer de una demanda que refleje de la forma más precisa la realidad (con la información del instante de llegada de los pasajeros y su paradero de destino) se define estimar los destinos de los pasajeros obtenidos en las etapas anteriores.

La estimación de destinos se realiza según una distribución gaussiana que tiene una media en la mitad de los paraderos posibles y una varianza = $\sqrt{2 * media}$.

En el caso que el algoritmo entregue un dato incoherente, se vuelve a repetir el procedimiento hasta obtener un destino dentro del rango de destinos factibles.

En la sección 2.2 siguiente se presentan las matrices origen-destino (OD), donde se aprecia el resultado de este algoritmo.

2.2. Análisis de demanda simulada

Cuando se procesan los datos facilitados por Transantiago S.A., se van teniendo pérdidas en las diferentes fases como se presentó en 2.1.1. Por esto, se tienen paraderos con información incompleta que no se pueden utilizar en ciertos horarios ya que se pueden distorsionar los análisis realizados.

Inicialmente se aplican dos criterios que permiten disponer de paraderos con la información más completa:

1. Se eliminan paraderos donde hubo cambio de recorrido, es decir, los paraderos donde el recorrido de los buses difiere de la trayectoria original (Se descartan los paraderos 14, 15, 16, 17, 18, 19, 20 y 74).
2. Se eliminan los que tienen una tasa de llegada promedio en algún día menor a un pasajero cada quince minutos (Se descartan los paraderos 1, 2, 26, 27, 28, 29, 42, 44, 45, 47, 52, 57, 62, 64, 66, 67, 70, 72, 73, 74, 76, 78, 79, 80, 82, 83, 84, 85).

Estos dos pasos anteriores permiten trabajar con 50 paraderos de los 85 paraderos iniciales. A pesar de ser una diferencia significativa se tiene información suficiente para simular un corredor que tiene paraderos con comportamiento real. Como se indicó previamente, se dispone de una semana completa de datos, pero estos siete días tienen comportamientos bien distintos, por lo que no se pueden utilizar todos con el objetivo de generar una demanda simulada de varios días para un día similar, por eso, a continuación se aplican ciertos criterios para seleccionar los días más similares.

Se analiza el volumen de pasajeros diarios de los 50 paraderos seleccionados, lo que se aprecia en la Figura 8:

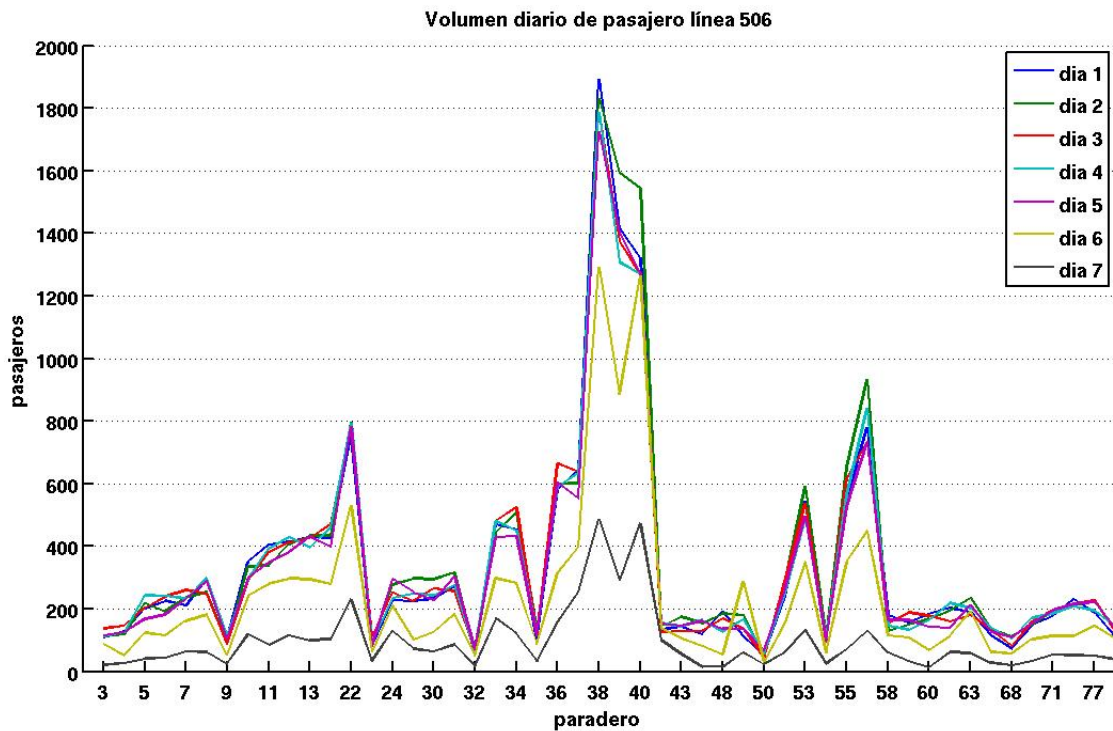


Figura 8: Volumen de paraderos

Al observar el volumen diario, se descartan los días 6 y 7 ya que tienen un comportamiento muy distinto a los otros 5 días.

Para implementar una estrategia de control para transporte público, se debe trabajar dentro de un rango horario que tenga un desarrollo similar, sobre todo que sea posible satisfacer la demanda sin cambios en la flota de buses. Por esto se escoje trabajar con el bloque de 7:00 a 10:00, ya que incluye la punta mañana en la demanda.

Se realiza nuevamente una selección de días según las curvas agregadas de pasajeros dentro de este bloque específico de tiempo. De este análisis se concluye descartar los días 1 y 5 por tener comportamientos distintos con respecto a los otros tres días. En la Figura 9 se presentan algunos paraderos.

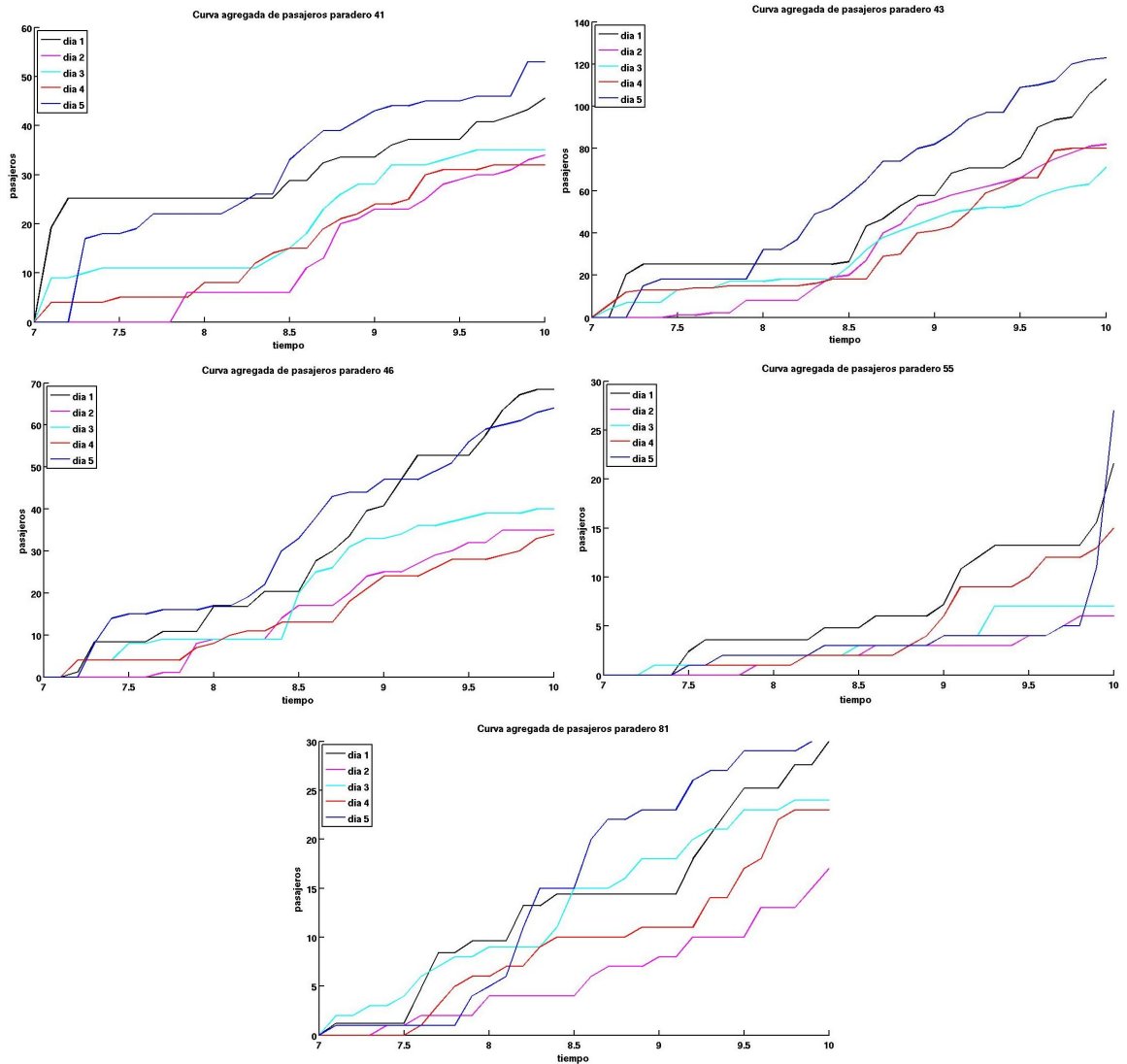


Figura 9: Curvas de carga agregada

Además se realiza nuevamente un análisis para descartar paraderos que tengan una alta diferencia entre los tres días que se utilizan. El procedimiento para calcular esta diferencia es considerar la máxima diferencia de cantidad de pasajeros en el bloque analizado entre los tres días y dividirlo por la cantidad máxima de pasajero de los tres días, de esa forma se obtiene una diferencia porcentual que se presenta en la tabla 6. De esta análisis se descartan los paraderos 9, 13, 32, 38, 39, 40, 41, 48, 65 por tener una diferencia mayor al 30 %.

Paradero	Día 2	Día 3	Día 4	Dif	Paradero	Día 2	Día 3	Día 4	Dif
3	81	93	102	21 %	39	31	45	35	31 %
4	160	161	171	6 %	40	17	19	25	32 %
5	113	161	145	30 %	41	31	35	46	33 %
6	90	95	121	26 %	43	82	71	80	13 %
7	148	148	167	11 %	46	35	40	34	15 %
8	89	92	83	10 %	48	29	20	22	31 %
9	34	44	49	31 %	49	108	142	144	25 %
10	90	97	118	24 %	50	127	168	144	24 %
11	171	148	140	18 %	51	85	75	72	15 %
12	22	50	54	59 %	53	32	45	33	29 %
13	42	34	29	31 %	54	19	27	20	30 %
21	484	525	470	10 %	55	13	12	15	20 %
22	36	51	37	29 %	56	55	68	48	29 %
23	46	47	39	17 %	58	32	40	31	23 %
24	76	75	85	12 %	59	191	252	192	24 %
25	100	115	100	13 %	60	30	21	24	30 %
30	57	61	46	25 %	61	31	28	35	20 %
31	191	209	186	11 %	63	653	560	489	25 %
32	40	45	31	31 %	65	39	31	64	52 %
33	74	74	63	15 %	68	29	39	41	29 %
34	92	73	76	21 %	69	107	89	125	29 %
35	77	66	55	29 %	71	110	77	89	30 %
36	78	104	82	25 %	75	52	46	42	19 %
37	315	334	304	9 %	77	58	41	41	29 %
38	13	20	29	55 %	81	17	24	23	29 %

Tabla 6: diferencia entre días disponibles para paraderos

Finalmente se seleccionan 40 paraderos de los 85 iniciales, lo cuales permiten generar una demanda que refleja de forma precisa la realidad de un posible sistema de transporte público. A cada pasajero se le estima un posible destino considerando lo descrito en 2.1.3, con lo que se obtiene como demanda representativa del sistema la matriz origen-destino presentada en la figura 10.

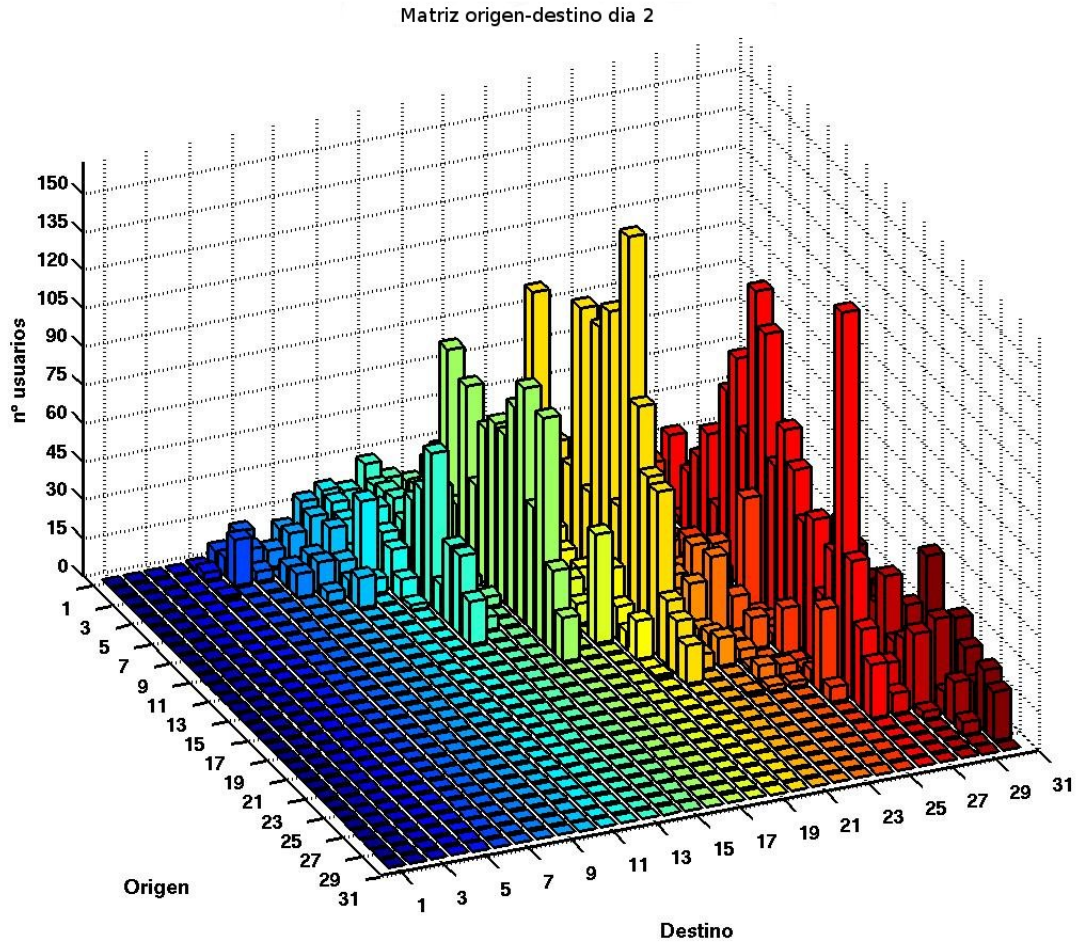


Figura 10: Matriz origen-destino día 2

2.3. Generación de nuevos datos simulados (réplicas)

Para poder realizar un análisis completo en una red de transporte público, se requieren más de tres réplicas de datos que son las disponibles en este momento gracias a lo facilitado por Transantiago S.A., entonces se propone el siguiente método para aumentar la cantidad de réplicas (o días) disponibles con lo que se pueden realizar análisis considerando disponibilidad de datos históricos.

- Para cada paraderos y día se modifica la cantidad de pasajeros hasta un máximo de 5 % según una variable aleatoria normal.
- Los instantes de llegada se modifican en forma aleatoria hasta un máximo de 5 %.

- Sobre este nuevo listado de instantes de llegada se aplica 2.1.3 para obtener los nuevos destinos.

Finalmente se tienen 30 días de réplicas (simulando la información disponible de un mes) para poder analizar diferentes estrategias de control con una confiabilidad suficiente.

2.4. Discusión

En este capítulo se presenta una estrategia para aprovechar los datos recolectados por un sistema de transporte público manejado por Transantiago S.A. que permite evitar realizar estudios específicos para determinar el tiempos de llegada, el origen y destino de pasajeros (con sus altos costos asociados), sino que se logra disponer de una demanda de pasajeros en paraderos confiable y que representa en forma precisa la realidad.

La estrategia presentada entrega una demanda de pasajeros que no se representa directamente a través de un modelo estadístico (como se ha hecho hasta el momento con el modelo de Poisson), por lo que la predicción de la misma es más compleja. Este punto es relevante para la aplicación de estrategias de control a sistemas de transporte público, ya que es necesario disponer de algoritmos de predicción precisos.

A continuación, en el capítulo 3, se presentan diferentes estrategias de control para sistemas de transporte público con la demanda propuesta.

3. Estrategia de control en tiempo real para sistemas de transporte público

3.1. Introducción

El diseño de sistemas de transporte público debe considerar una serie de variables a determinar tales como el número de líneas y sus rutas asociadas las cuales están fuertemente relacionadas con la intensidad de demanda. Además el diseño debe ser realizado considerando diferencias entre periodos establecidos (horas puntas, valle) para lograr un ajuste adecuado, Ortúzar (1983). En situaciones reales, el movimiento de los buses en sistemas de transporte público se ve afectado por diferentes interrupciones: tráfico de congestión, desfases inesperados, aleatoriedad de la demanda de pasajeros (tanto temporal como espacial) accidentes y así sucesivamente. Estos eventos dificultan el despacho de buses al igual que la operación de los buses en ruta cuando siguen un esquema preplanificado definido sobre un promedio en cada periodo de operación. Entonces para lograr disminuir el efecto de estas perturbaciones se comienzan a desarrollar estrategias de control flexible tanto en tiempo real como fuera de línea. La evolución histórica de estas estrategias va desde la asignación de flotas pre-planificadas y estrategias de programación hasta análisis de estrategias de control en tiempo real, considerando información en tiempo real disponible tal como conteo automático de pasajeros (*Automatic Passenger Counters*, APC) y localización automática de vehículos (*Automatic Vehicle Location*, AVL). Se han diseñado un grupo de estrategias para permitir al operador reaccionar dinámicamente a las distorsiones del sistema en tiempo real. La estrategia más estudiada ha sido *holding* orientada en la mayoría de los casos a mantener un *headway*² entre los buses lo más cercano posible de un valor predefinido. Además existen otros tipos de estrategias de control como lo es el *station skipping* (*expressing*), Suh et al. (2002), Fu & Liu (2003).

En términos de la configuración espacial de diferentes estrategias de control, Eberlein (1995) las clasificó en tres categorías: control en paraderos, control entre paraderos y otras

²Diferencia de tiempo entre dos buses viajando en la misma dirección sobre la misma ruta

estrategias. Las estrategias de control en paraderos son de dos tipos: *holding* y *station skipping* (*deadhealing*, *expressing*, *short-turnig*, etc). Estrategias de control entre paraderos incluye control de velocidad y prioridad por señal de tránsito, entre otras. Otras estrategias incluye por ejemplo, división de tren, la cual es más orientada a control de sistemas de trenes en la literatura.

Con respecto a las contribuciones más notables en el estudio de estrategias de *holding*, se pueden mencionar a Barnett (1974), Turnquist & Blume (1980), Eberlein (1995), Eberlein et al. (2001), Hickman (2001), Sun & Hickman (2004), Zolfaghari et al. (2004) y Yu & Yang (2007).

Barnett (1974) desarrolló un modelo simple de *holding* en una estación de control, donde la suma de los tiempos totales más el tiempo extra de espera de los pasajeros a bordo es minimizado. Turnquist & Blume (1980) identificaron las condiciones bajo las cuales los resultados de *holding* son atractivos. El estudio de Hickman (2001) presentó un modelo de *holding* estocástico en una estación de control dada. Investigaciones más recientes han explorado modelos de *holding* que se basan en información en tiempo real, principalmente referidas a la ubicación de vehículos (Eberlein (1995), Eberlein et al. (2001), Hickman (2001), Sun & Hickman (2004)).

Eberlein (1995), Eberlein et al. (1999) y Eberlein et al. (2001) establecieron una programación cuadrática determinista bajo planificación de horizonte móvil, en el cual la decisión de *holding* para un vehículo específico afecta la operación de un subconjunto de vehículos precedentes. Se concluye que tener más de dos estaciones de *holding* en un corredor no es necesario. Estos resultados contradicen a los de Sun & Hickman (2004). Su trabajo concluye que el *holding* múltiple de vehículos en múltiples estaciones produciría un mejor resultado que en una única estación. La mayoría de estos modelos propone heurística para resolver los problemas en vez de complejas formulaciones matemáticas. Zolfaghari et al. (2004) desarrollaron un modelo de control matemático para *holding* usando información en tiempo real de la ubicación de los buses. En este trabajo, los tiempos de espera son calculados a partir de la diferencia de tiempos de salida de los buses y la función objetivo minimiza los tiempos de espera de los

usuarios que arriban a una parada y aquellos que tienen que esperar más de un bus.

Finalmente, Yu & Yang (2007) presentan una estrategia de *holding* dinámico, en la cual, se calcula el desempeño on-line de la operación temprana de los buses para la siguiente parada y se optimiza el tiempo de *holding*. También se desarrolla un modelo basado en máquina de soporte de vectores (*Vector support machine*, VSM) para la predicción de tiempos de salida de buses desde la siguiente parada.

En relación a las estrategias de control *station skipping*, éstas han sido estudiadas considerando su aplicación en forma predefinida (Jordan & Turnquist (1979); Furth (1986)) y, de forma más extensa, como estrategias de control en tiempo real (Lin et al. (1995); Eberlein (1995); Eberlein et al. (1999); Fu & Liu (2003); Sun & Hickman (2005)). En el último caso, el enfoque consiste en acelerar los buses saltando estaciones (una o más) para recuperar su itinerario pre-planificado después de interrupciones o inesperados atrasos, con el fin de reducir el impacto en el nivel de servicio medido a través de tiempos de espera de usuarios en paraderos más el tiempo extra de espera de las estaciones saltadas. En general una acción de *station skipping* es definida antes de la partida de los buses del terminal, excepto en el modelo de Sun & Hickman (2005), que permite tomar acciones de control una vez que el vehículo está en la ruta.

Eberlein (1995) formuló un modelo integrado, que abarca *holding*, *deadhealing* y *expressing*. Adicionalmente, Adamski & Turnau (1998) y Adamski (1996) desarrollaron una herramienta de soporte de decisiones simuladas para despacho dinámico óptimo, incluyendo control de puntualidad (donde se compensan desviaciones desde la planificación), control de regularidad (donde se compensan desviaciones desde el *headway* regular) y control sincronizado basado en control de retroalimentación lineal cuadrático mientras se consideran las restricciones de estado del sistema.

En esta tesis se utiliza el modelo presentado en Sáez et al. (2009) que desarrolla una propuesta integrando dos de las estrategias nombradas (*holding* y *station skipping*), para resolver un problema de transporte público en tiempo real con incerteza en la demanda de pasajeros, considerando información online del comportamiento del sistema. El modelo de decisión de

estaciones *skipping* se realiza de una manera similar a Sun & Hickman (2005), es decir, en tiempo real, lo cual permite que la estructura propuesta sea más adaptable y sensible a atrasos en tiempo real.

El modelo es formulado como un problema de control predictivo híbrido (HPC), en que la teoría fundamental se ajusta a las condiciones dinámicas típicas del transporte público que disponen de variables continuas y discretas o cuantizadas. Los modelos predictivos permiten la estimación de efectos futuros de las acciones de control en el comportamiento del sistema de buses y también permiten la inclusión de complejos sistemas de restricciones. Ellos también tienen la capacidad de optimizar el desempeño del sistema en tiempo real basado en la elección de una apropiada función objetivo (Hegyi (2004); Hegyi et al. (2005); Karer et al. [2007a,2007b]). Aún más, los enfoques predictivos son adecuados para ambientes dinámicos con alta incerteza en el futuro de los eventos, los cuales pueden hacerse relevantes para el proceso de toma de decisiones que tiene que ser desarrollado en tiempo real.

Para este problema el controlador predictivo corresponde al despachador de buses, quien dinámicamente entrega acciones de control óptimas al sistema de buses para optimizar el desempeño de acuerdo a una función objetivo asumida que toma en cuenta la futura evolución del sistema de transporte público. En este caso particular, la función objetivo dinámica está mayormente orientada a minimizar los tiempos totales de viaje de los usuarios (incluyendo tiempos de espera como también tiempos sobre el vehículo).

La demanda de pasajeros en tiempo real, la cual es desconocida e incierta, es modelada como una perturbación para el esquema predictivo, dado que diferentes patrones de llegadas de pasajeros podrían afectar significativamente el tiempo estimado de viaje dentro de los buses para operaciones de transferencia más largas en los paraderos. Las estrategias de control permiten incorporar en el diseño el futuro comportamiento del sistema total asociado con la operación de los buses, usando un sistema de predicción para las distorsiones (demanda). La metodología para la predicción de demanda necesita una buena estimación de las matrices origen destino en tiempo real para apropiadamente predecir el número futuro de pasajeros en paraderos, carga de buses y tiempos de transferencia de pasajeros, basados en datos en

tiempo real como históricos.

La función objetivo del modelo considerado emplea, entre otras, una componente relacionada con la regularización del *headway*, basándose en varios autores en la literatura de control en tiempo real de transporte público (Welding (1957); Osuna & Newell (1972)) que la han utilizado en el diseño de las estrategias ,ya que esta componente está estrechamente relacionada con el objetivo de minimización del tiempo total de espera porque éste se incrementa en forma lineal con la variación del *headway*.

A continuación se presenta el detalle del modelo utilizado.

3.2. Formulación del problema

El sistema utilizado en el modelo es un corredor circular simple (aunque su implementación a sistemas más complejos es posible) de una vía con alta demanda de pasajeros con P paraderos operados por un servicio simple de alta frecuencia consistente en b buses homogéneos, con una capacidad fija denotada \bar{L} . Los vehículos comienzan su recorrido en el terminal definido como paradero 1 visitando todos los paraderos siguientes $(2, 3, \dots, P)$ antes de retornar al mismo paradero terminal (1) donde todos los pasajeros deben descender.

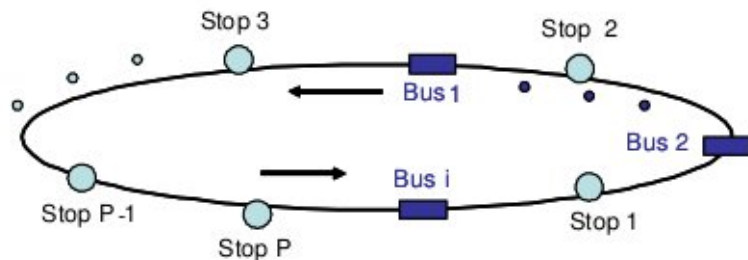


Figura 11: Modelo de corredor circular

El sistema propuesto incorpora las siguientes suposiciones:

- Los buses sirven todas las paradas y el adelantamiento no está permitido.
- Las tasas de llegada de pasajeros a los paraderos están definidas basándose en la demanda de Transantiago S.A. (según Capítulo 2).

- En el sistema, la llegada de los pasajeros a cada estación se considera desconocida y la distribución de los posibles destinos también. Cada pasajero se identifica por el paradero de llegada y destino, además del instante de arribo al paradero. El impacto de las formas de distribución de demanda en el algoritmo de control se analiza con mayor detalle en el capítulo 4.
- La propuesta considera una doble dimensionalidad del modelo dinámico: espacial y temporal con características continuas (posición de los buses y velocidad) y discretas (número de pasajeros en los buses), por lo que un enfoque de sistema híbrido permite representar ambos tipos de variables (Bemporad & Morari (1999)). Los eventos, identificados por k , son producidos por una acción específica, que en este sistema, corresponden a las llegadas de un bus a una estación por lo que el paso de discretización es variable. El tiempo continuo se identifica con t , y t_k corresponde al instante en que ocurre el evento k . Se define trabajar con las llegadas de los buses a paraderos porque es el instante en que se pueden aplicar acciones de control al sistema. Considerar que un evento k está siempre asociado con la llegada de un bus i específico a una parada p específica.

El diagrama de bloques del sistema de control propuesto en Sáez et al. (2009) se presenta en la Figura 12.

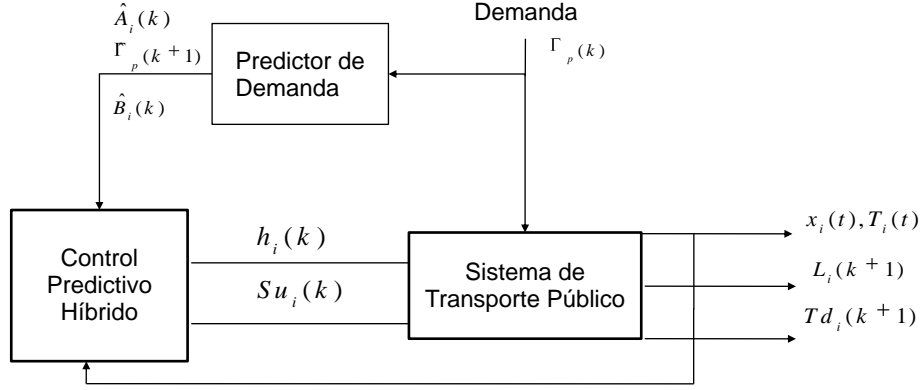


Figura 12: Diagrama de estrategia control predictivo

Para cada bus i de la flota, se definen su posición, $x_i(t)$, en cualquier instante t continuo y el tiempo que falta para que el bus i alcance la siguiente parada, $T_i(t)$, con el objetivo de chequear el estado de los buses y consecuentemente generar los eventos. Un nuevo evento k es generado por el bus i en cualquier parada p cuando $x_i(t)$ alcanza la posición de esta parada en $t = t_k$. Por consiguiente, el tiempo que falta para que el bus i alcance esta parada es igual a cero ($T_i(t_k) = 0$).

En la Figura 12 las variables manipuladas corresponden a *holding* $h_i(k)$ y a *station skipping* $Su_i(k)$ que se asocian a un bus i y a un evento k . $h_i(k)$ es el lapso durante el cual el bus i se retiene en la parada asociada al evento k . $Su_i(k)$ es una variable binaria que es igual a uno si se permiten a los pasajeros abordar al bus i en el paradero asociado al evento k , cero en otro caso. Analíticamente se definen a continuación

$$Su_i(k) = \begin{cases} 1 & \text{con } \Theta(i, k) \\ 0 & \text{en caso contrario} \end{cases} \quad (1)$$

Donde la condición $\Theta(i, k)$ es verdadera si a cualquiera de los pasajeros se les permite abordar

el bus i o si cualquiera de los pasajeros arriba del bus alcanza su destino en el evento k .

$$h_i(k) = n_i\tau, \quad n_i \in \mathbb{Z}^+, \tau > 0 \quad (2)$$

Esta expresión significa que los periodos de *holding* son múltiplos de un paso fijo τ . Se utiliza esta simplificación para facilitar la formulación y aplicación de la solución del algoritmo.

Las salidas en tiempo discreto corresponden a la carga de pasajeros $L_i(k+1)$ y al tiempo de partida $Td_i(k+1)$ una vez que el bus parte desde la actual parada, asociado con el bus i que generó el evento k .

$\Gamma_p(k)$ es el número de pasajeros esperando por un bus en el paradero p y se modela como una perturbación por medio del estimador de demanda; las variables $\hat{A}_i(k)$, $\hat{B}_i(k)$ y $\hat{\Gamma}(k+1)$ son estimadas e incorporadas al modelo dinámico.

Estas variables se definen como:

$\hat{A}_i(k)$: Representa la cantidad de pasajeros que bajan del bus i en el instante k .

$\hat{B}_i(k)$: Es el número esperado de pasajeros que abordarán el bus i durante cierto intervalo.

$\hat{\Gamma}_p(k+1)$: Cantidad de pasajeros estimada que se encuentra esperando después de t_k en el paradero p cuando pasa el bus i asociado al evento k .

A continuación se presenta el modelo de estrategia de control para el problema presentado.

3.3. Estrategia de control

El modelo predictivo describe el comportamiento dinámico de las principales variables indicadas anteriormente, como función de las acciones de control.

Primero, la posición esperada de un bus en el instante t , $\hat{x}_i(t)$, es descrita como una función de la velocidad instantánea del bus $v_i(t)$ que depende en el tiempo continuo y de las acciones de control aplicadas. Se comienza calculando la posición del bus i en el tiempo continuo t como sigue:

$$\hat{x}_i(t) = x_i(t_k) + \int_{t_k}^t \hat{v}_i(\partial) d\partial \quad (3)$$

donde t_k es el instante continuo en el cual el evento k es generado y $x_i(t_k)$ la posición del bus i en el instante t_k . La velocidad instantánea se modela asumiendo una velocidad constante v_0 cuando el vehículo se encuentra en movimiento y cero en otro caso, lo cual implica que el proceso de aceleración y desaceleración es ignorado. La Figura 13 muestra la función velocidad del bus i mientras viaja desde la estación que alcanza en el instante k hasta que el bus arriba a la siguiente parada (que es asociado con el instante futuro $k + d$). Notar que d corresponde a los instantes de tiempo generados por otros buses de la ruta llegando a diferentes paradas, producidos mientras el bus i viajaba entre la actual parada y la siguiente. En la Figura 13, $\hat{T}r_i(k)$ es el tiempo estimado, asociado con la transferencia de pasajeros (máximo entre subida y bajada) y $\hat{T}v_i(k)$ es el tiempo estimado de viaje entre las dos estaciones consecutivas. Como se ha indicado anteriormente, el controlador decide el tiempo de *holding* en la estación p denominado $h_i(k)$. Claramente, cuando un bus se encuentra en la parada la velocidad es igual a cero, mientras el bus se encuentra transfiriendo pasajeros también y además durante el periodo de *holding*, lo cual significa que la velocidad instantánea actualmente depende de esas variables.

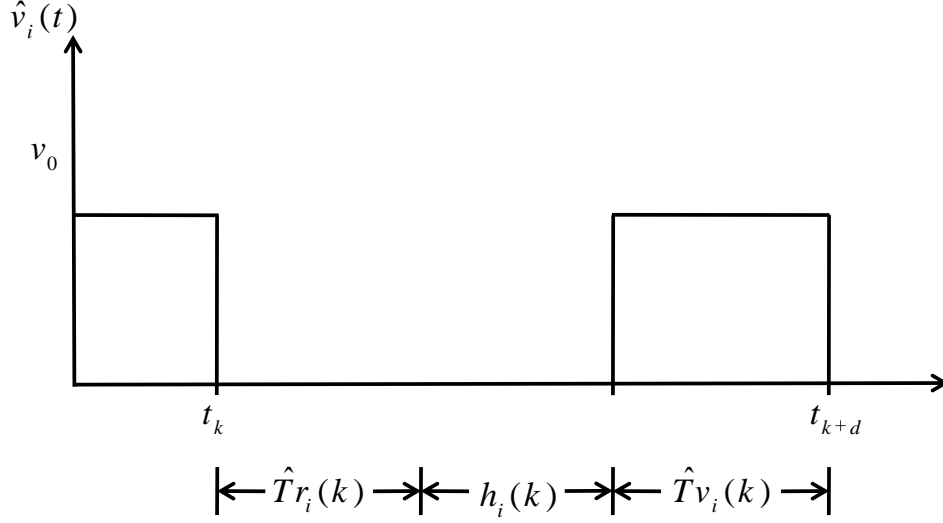


Figura 13: Ejemplo de velocidad de bus entre paradas consecutivas

En este contexto y basado en la Figura 13, una estimación de la velocidad instantánea puede ser calculada como:

$$\hat{v}_i(t) = \begin{cases} 0 & t_k \leq t \leq t_k + \hat{T}r_i(k) + h_i(k) \\ v_0 & t_k + \hat{T}r_i(k) + h_i(k) \leq t \leq t_{k+d} \end{cases} \quad (4)$$

Para generar el siguiente evento del modelo dinámico, se requiere el tiempo esperado que le queda al bus i para alcanzar la siguiente parada (medido desde el instante t); éste se puede calcular como:

$$\hat{T}_i(t) = t_k + Su_i(k) \cdot h_i(k) + \hat{T}r_i(k) + \hat{T}v_i(k) - t, \quad t_k < t \leq t_{k+d} \quad (5)$$

A continuación se define y se calcula analíticamente la predicción de las variables de salida discretas del modelo dinámico, requerido por la estrategia HPC; $\hat{L}_i(k+1)$ y $\hat{T}d_i(k+1)$.

Primero, se debe definir la carga predicha de pasajeros, $\hat{L}_i(k+1)$, como el número estimado de pasajeros en el bus i una vez que parte de la estación, analíticamente:

$$\hat{L}_i(k+1) = \begin{cases} \min [\bar{L}, L_i(k) + Su_i(k) \cdot \hat{B}_i(k) - \hat{A}_i(k)] & \text{si el bus } i \text{ genera el evento } k \\ L_i(k) & \text{otro caso} \end{cases} \quad (6)$$

Donde \bar{L} es la capacidad del bus, $L_i(k)$ es la carga del bus en el instante k , $\hat{B}_i(k)$ corresponde al número esperado de pasajeros que abordarán el bus i y $\hat{A}_i(k)$ representa el número estimado de pasajeros que descienden del bus i en el evento k .

Notar que $\hat{A}_i(k)$ y $\hat{B}_i(k)$ se obtienen a través de análisis estadístico de la recolección de datos desde sensores que deberían ser localizados en los paraderos y buses. En este enfoque, estas estimaciones se obtienen desde datos que provienen desde días similares previos (información histórica registrada) e información dinámica que ocurre el mismo día (información en tiempo real).

Basado en la información histórica registrada, es posible estimar $\hat{A}_i(k)$ usando el patrón de destino más frecuente desde los días previos sobre el mismo periodo de tiempo; luego, estas estimaciones se corrigen con la información en tiempo real obtenido desde las preferencias observadas en pasajeros que se encuentran en el sistema. $\hat{B}_i(k)$ se calcula en base a la carga estimada del bus $\hat{\Gamma}_p(k)$ en el instante k y la capacidad del bus. $\hat{\Gamma}_p(k)$ es estimado considerando un modelo sobre promedios autoregresivo móvil para el tiempo de llegada de los pasajeros en las paradas. Por otra parte, el tiempo estimado de transferencia definido anteriormente puede ser analíticamente calculado como $\hat{T}r_i(k) = \text{Max} (t_a \cdot \hat{A}_i(k), t_b \cdot \hat{B}_i(k))$ donde t_a y t_b son las tasas marginales de subidas y bajada respectivamente.

En adición, el tiempo de partida del bus i $\hat{T}d_i(k+1)$ puede ser calculado como:

$$\hat{T}d_i(k+1) = \begin{cases} t_k + Su_i(k) \cdot h_i(k) + \hat{T}r_i(k) & \text{si el bus } i \text{ genera el evento } k \\ Td_i(k) & \text{otro caso} \end{cases} \quad (7)$$

La predicción de la carga del paradero $\hat{\Gamma}_p(k+1)$, definido como el número de pasajeros esperando en la estación p asociado con el bus i que genera el evento k , se puede calcular

como:

$$\hat{\Gamma}_p(k+1) = \begin{cases} \Gamma_p(k) + \hat{\delta}_p(k) - B_i(k) & \text{si el bus } i \text{ genera el evento } k \\ \Gamma_p(k) + \hat{\delta}_p(k) & \text{otro caso} \end{cases} \quad (8)$$

Donde $\Gamma_p(k)$ corresponde a la carga del evento anterior en el mismo paradero p , $\hat{\delta}_p(k)$ es la cantidad de pasajeros que llega a esta estación entre el evento anterior k y $k+1$, se genera basándose en análisis estadístico de los datos en los días similares previos y el mismo día (datos históricos y en tiempo real). Los datos históricos consideran la medición del origen destino de los pasajeros de cada paradero durante un periodo de tiempo (ejemplo, 2 semanas previas). Los datos en tiempo real se utilizan para corregir la estimación obtenida usando datos históricos. El detalle de este algoritmo de predicción se presenta en la sección 3.6. Las variables relacionadas con la demanda $\hat{A}_i(k)$, $\hat{B}_i(k)$, $\hat{\Gamma}_p(k+1)$ son de gran relevancia para el sistema ya que dimensionan los requisitos que se deben satisfacer, pero al mismo tiempo son de difícil estimación y predicción, por no disponer de un patrón claro y ser un sistema no lineal, por lo cual el *predictor de la demanda* posee una alta importancia dentro del diagrama de estrategia de control, Figura 12. Para la predicción de las variables relacionadas con la demanda se propone una nueva metodología basada en redes neuronales, lo que se presenta en el Capítulo 5.

Utilizando la ecuación (7) sobre el tiempo de partida, es posible predecir el *headway* $\hat{H}_i(k+1)$ del bus i que genera el evento k , con respecto al bus precedente $i-1$ cuando alcanzan la misma parada, lo cual corresponde al evento $k+1-z_{i-1}$. Analíticamente:

$$\hat{H}_i(k+1) = \hat{T}d_i(k+1) - \hat{T}d_{i-1}(k+1-z_{i-1})$$

donde $\hat{T}d_i(k+1)$ es asociado con el bus i que genera el evento k , y $\hat{T}d_{i-1}(k+1-z_{i-1})$ representa el tiempo de partida predicho del bus precedente $i-1$ que genera el evento $k-z_{i-1}$, en la misma parada. La variable z_{i-1} representa el número de eventos entre la llegada del bus precedente $i-1$ y el bus i , ambos alcanzando la misma parada.

3.4. Restricciones operacionales

El modelo predictivo del sistema de transporte público debe satisfacer algunas restricciones físicas y operacionales.

La primera restricción corresponde a la restricción de capacidad (ya detallada en la ecuación (6)). Esta es una restricción física en el sentido que el bus no puede transportar más pasajeros que la capacidad máxima.

La restricción de precedencia y la consistencia de la demanda, ambas son relevantes, porque cada pasajero tiene un origen y destino específico. La restricción de precedencia evita que pasajeros puedan bajarse antes de que ellos tomen el bus. En relación a la demanda, se asume que no hay nodos de transferencia, y por ende, una vez que los pasajeros se encuentran dentro del bus, se bajarán desde el mismo bus en su destino. También una vez que el pasajero arriba a su destino, siempre se baja del bus ahí.

Considerando la operación de los buses, el modelo restringe a los buses a parar en un paradero si cualquier pasajero lo requiere, es decir, si ese paradero es el destino de algún pasajero, inclusive si el modelo recomienda aplicar la estrategia de control *station skipping*, de forma similar como se aplica en Sun & Hickman (2005). Esta restricción se aprecia que trabaja mejor que incluir alguna penalización en la función objetivo, en cuyo caso, algún pasajero podría terminar bajándose en una estación diferente a la de su destino, lo cual no es deseable. Por otro lado, si el modelo recomienda aplicar la estrategia de control *holding* en cierto paradero, lo cual no es físicamente apropiado como operación, se tiene que el bus sólo parará en la parada durante el lapso requerido para la transferencia normal de pasajeros, es decir, considera tiempo de *holding* cero.

En relación a las acciones de control, tanto *holding* como *skipping* pueden ser aplicadas en todas las paradas.

Cada bus es identificado por una etiqueta única. De todas formas, el modelo permite actualizar los índices cuando un bus llega a la siguiente parada, ordenados de tal forma que el bus $i - 1$ siempre precede al bus i .

3.5. Función objetivo

El paso siguiente es definir en forma apropiada la función objetivo, para poder tomar decisiones en tiempo real y optimizar la dinámica del sistema.

- Función para minimizar los tiempos de espera de los pasajeros

$$J_1 = \sum_{l=1}^{N_p} [\theta_1 \cdot \hat{H}_i(k+l) \cdot \hat{\Gamma}_p(k+l)] |_{i=i(k+l-1)} \quad (9)$$

Donde $\hat{H}(k+l)$ corresponde al tiempo que transcurre entre la salida de un bus de un paradero en un instante futuro $k+l$ y la salida de su predecesor. θ_1 es el factor de ponderación para sintonizar y ajustar la importancia que se le quiera dar al término.

Este término considera la minimización del tiempo de espera dentro de un horizonte de predicción N_p para cada bus i que genere un evento en $k+l$.

- Función para ajustar la regulación al *headway* de referencia

$$J_2 = \sum_{l=1}^{N_p} [\theta_2 \cdot (\hat{H}_i(k+l) - \bar{H})^2] |_{i=i(k+l-1)} \quad (10)$$

Este término está asociado a la regularización de intervalos, minimizando la diferencia entre intervalos en torno a un intervalo óptimo, donde $\hat{H}(k+l)$ corresponde al tiempo que transcurre entre la salida de un bus en un instante futuro $k+l$ y la salida de su predecesor, \bar{H} corresponde al intervalo de diseño que satisface la demanda de la red durante cierto periodo. Normalmente, el diseño del *headway* está relacionado con el diseño de la frecuencia que depende directamente de la carga de los segmentos y puede ser determinado como el mínimo requerido para movilizar los pasajeros del segmento más cargado a través de la ruta. En sistemas más sofisticados, el diseño de la frecuencia es calculada para minimizar la función objetivo estática que involucra los costos del operador como del usuario, en cuyo caso, la frecuencia óptima debería ser mayor que la frecuencia mínima que es capaz de movilizar todos los pasajeros en un nivel agregado. θ_2 es el factor de ponderación para sintonizar y ajustar la importancia que se le quiera dar al término. Esta componente es importante porque permite disminuir la variabilidad de la regulación de intervalos, acercándolos a un valor de referencia de diseño.

- Función de minimización asociado al *holding*

$$J_3 = \sum_{l=1}^{Np} [\theta_3 \cdot \hat{L}_i(k+l) \cdot h_i(k+l-1)] \Big|_{i=i(k+l-1)} \quad (11)$$

Considera el tiempo extra de espera de los pasajeros que están arriba del bus que realiza *holding*. θ_3 es el factor de ponderación para sintonizar y ajustar la importancia que se le quiera dar al término.

- Función de minimización asociado a *station skipping*

$$J_4 = \sum_{l=1}^{Np} [\theta_4 \cdot \hat{\Gamma}_p(k+l) \cdot \hat{H}_{i+1}(k+l+z_{i+1}) \cdot (1 - Su_i(k+l-1))] \Big|_{i=i(k+l-1)} \quad (12)$$

El *station skipping* agrega un tiempo de espera considerable para la gente que se encuentra en esa estación. Si se realiza esta acción de control $Su_i(k+l-1)$ toma el valor 1. La acción debe ser realizada en el instante anterior $k+l-1$ y si no se realiza la acción, se anula el término. θ_4 es el factor de ponderación para sintonizar y ajustar la importancia que se le quiera dar al término.

Por lo tanto, la función a minimizar es

$$\underset{\{u(k), \dots, u(k+Np-1)\}}{\text{Min}} J = J_1 + J_2 + J_3 + J_4 \quad (13)$$

donde $u(k), \dots, u(k+Np-1)$ es la secuencia de acciones de control con $u(k+l-1) = \begin{bmatrix} Su_i(k+l-1) \\ h_i(k+l-1) \end{bmatrix}^T$ si el bus i genera el evento $k+l-1$.

Considerar que el propósito de la función objetivo es satisfacer las expectativas de los usuarios en relación a los tiempos de espera y viaje, propósito que se aborda a través de un esquema a nivel operacional y no un esquema a nivel de optimización de inversión o costos de operación. Por esto, se asume un tamaño de flota fija obtenido desde el diseño de frecuencia necesaria para satisfacer la demanda de pasajeros, y el beneficio que se busca con la aplicación de las

estrategias de control es la calidad de servicio de los pasajeros. Dadas estas consideraciones, las componentes de costo no se consideran en la función objetivo, aunque en otras condiciones, éstas deberían convertirse en una componente importante en las decisiones en tiempo real para no sub-utilizar los recursos asignados al corredor.

3.6. Modelo de predicción de la demanda

Para poder desarrollar el modelo HPC se requiere de un bloque que permita predecir las perturbaciones, en este caso la demanda de pasajeros, y de esa forma poder escoger la mejor acción de control, según cómo se comporte el sistema en el futuro.

Los datos asociados a los pasajeros que se predecirán se caracterizan por el paradero de origen y el paradero de destino, además del instante de llegada al paradero de origen. Estos datos permiten obtener una matriz origen destino representativa de la demanda para cada periodo de modelación.

El modelo presentado en Sáez (2007), conocido como *modelo de predicción heurístico basado en promedio*, consiste en predecir los instantes de llegada de los siguientes pasajeros a partir de un tiempo t_k de forma recursiva para obtener el arribo de todos los pasajeros dentro del intervalo de análisis comprendido entre dos eventos consecutivos de llegada de buses a un mismo paradero $\hat{\Delta}_p(k)$. Para lograr lo anterior, se asume que se cuenta con información de días pasados y en tiempo real de la cantidad de usuarios que arriban al paradero y el destino de cada uno.

La estimación se realiza con una predicción simplificada en ventanas móviles: la hora de llegada de una persona a un paradero se realiza mediante promedios ponderados; mientras el destino de esta persona corresponde al destino más frecuente de la ventana generada, el algoritmo se detalla en los anexos.

Se debe aclarar que el algoritmo de control no necesita específicamente los instantes de llegada de los pasajeros, sino solamente la cantidad de pasajeros que llegaron entre dos eventos, necesidad que es abordada en el *modelo de predicción heurístico basado en promedio* a través de la predicción específica de instantes de llegada de pasajeros a paraderos.

Este modelo de predicción posee una alta precisión si la demanda tiene tasas de llegada más o menos constantes, ya que como está basado en promedios por lo que tiende a replicarlas, pero si se tienen modelos con distribuciones aleatorias, el resultado es menos exacto. Por esto se propone un nuevo algoritmo de predicción basado en redes neuronales (capítulo 5).

3.7. Método de solución del controlador predictivo híbrido basado en estrategia de todos los paraderos

Una forma para resolver la función objetivo de optimización es hacerlo a través de numeración explícita de todas las soluciones. Este es el enfoque más simple y directo, ya que considera analizar todas las opciones existentes.

Comienza tomando en cuenta el sistema actual real con las posibles acciones para el siguiente evento y va simulando pasos hacia futuro partiendo de todas las posibles alternativas, lo que genera un árbol con las acciones de control factibles. No siempre es posible aplicar todas las acciones de control diseñadas en una parada ya que, por ejemplo, en el paradero donde se genera el evento k algún pasajero del bus i puede desear bajarse por lo que no se podría realizar *skipping*.

El controlador predictivo va considerando cada una de estas acciones, lo cual se hace sucesivamente hasta los Np pasos futuros del controlador. Esto implica que se tendrán que analizar, en el caso más completo, $[acciones\ control]^{Np}$.

La ventaja de este sistema es que entrega en forma precisa la mejor acción de control para el siguiente evento que tiene el mejor comportamiento a Np pasos futuros, pero tiene ciertas desventajas que se indican a continuación:

- Tiempo de cálculo extenso: si se desea mejorar la acción de control escogida, es necesario tener la mayor cantidad de pasos futuros Np ya que esto permite observar de mejor forma las consecuencias de las decisiones tomadas en el instante actual, pero eso implica aumentar en forma considerable las opciones a analizar (aumenta en forma geométrica) y por ende, el tiempo de cálculo.

- En relación a lo anterior, en sistemas con una cantidad considerable de paraderos y buses, se necesita tener un número importante de pasos futuros Np . porque si la red es de gran tamaño, lo que sucede en un paradero no tiene un impacto considerable en el corto plazo en los otros paraderos más alejados, entonces, si Np es pequeño comparado con la cantidad de buses y paraderos, los eventos predichos se distribuyen en forma más o menos homogénea en los diferentes paraderos de la red, por lo que no se logra visualizar el impacto local hacia futuro de la acción analizada. Esto se puede enfrentar de dos formas, una es aumentando considerablemente Np (pero se tiene como consecuencia la primera desventaja) o lo otro es generando una predicción local de eventos.

En la sección 3.8 se presenta un nuevo método de solución que considera una predicción local de eventos para poder enfrentar las desventajas del método recién evaluado.

3.8. Método de solución del controlador predictivo híbrido basado en estrategia de selección de paraderos

Este algoritmo de solución tiene su origen en la mejora de una de las desventajas del método de solución basado en estrategia de todos los paraderos mencionado anteriormente. Si el objetivo es evaluar el enfoque en un sistema de tránsito realista, con muchos buses y paradas, la probabilidad de que el próximo evento futuro suceda cerca de una parada específica es baja (ejemplo, el paradero donde el evento actual ocurre); entonces si se predicen n pasos hacia adelante, los futuros eventos que realmente afecten a la actual decisión serán muy pocos, reduciendo el poder predictivo del enfoque. La premisa es que los eventos analizados en el futuro se distribuirán a través de todo el sistema, lo que no permite enfocarse en el análisis del impacto real de la decisión actual de la parada dada. Este problema se podría evitar incrementando los pasos futuros n considerablemente, para asegurar que se tiene una cantidad suficiente de eventos que realmente se afectan por la decisión que se debe tomar, lo cual se presentó en la sección 3.7, de todas formas, este enfoque no es computacionalmente factible para implementaciones reales.

Para evitar este problema, se propone implementar una estrategia que permita un análisis

local del impacto de las acciones de control para poder compararlas, evitando una detallada figura de las acciones que ocurren lejos del paradero donde se debe tomar la decisión. De esta forma el controlador sólo debe predecir los eventos que suceden en una vecindad P^* del paradero, donde el evento actual ocurre. Como consecuencia, se disocian los pasos futuros del controlador de los pasos futuros del simulador. Luego, entre la ocurrencia de dos eventos dentro de la vecindad P^* , varios otros eventos pueden ocurrir fuera de los límites, lo cuales no son realmente significantes en la selección de la mejor acción a aplicar. El estado de los buses fuera de la vecindad P^* se aproxima evitando que los buses se detengan en los paraderos, por lo que se debe tener en cuenta que mientras más pequeña es la vecindad P^* mayor es el error de la posición de los buses fuera de ésta. Por eso se debe equilibrar adecuadamente la cantidad de pasos futuros a predecir y el tamaño de P^* .

Para implementar este método, solamente se debe realizar una modificación en el condicionamiento del *skipping*. Recordando la ecuación (1):

$$Su_i(k) = \begin{cases} 1 & \text{if } \Theta(i, k) \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

La modificación va en la condición $\Theta(i, k)$ que originalmente es verdadera si a cualquiera de los pasajeros se les permite abordar el bus i o si cualquiera de los pasajeros en el bus alcanza su destino en el evento k . A las restricciones anteriores se le superpone la condición de que el paradero $p(k)$ donde se genera el evento esté dentro de P^* para que $\Theta(i, k)$ sea verdadero.

3.9. Análisis de sensibilidad de métodos de solución

Los dos métodos de solución: método de solución del controlador predictivo híbrido basado en estrategia de todos los paraderos (T_parad), sección 3.7 y método de solución del controlador predictivo híbrido basado en estrategia de selección de paraderos (S_parad), sección 3.8, tienen varios parámetros que son posibles de ajustar para obtener mejores resultados de predicción, por lo que se realiza un análisis de sensibilidad para observar el comportamiento de los predictores al modificar sus parámetros.

Este análisis se debe desarrollar basado en ciertos indicadores que permitan reflejar el buen o mal comportamiento del controlador según el objetivo de éste, por esto se escogen los siguientes indicadores:

- Diferencia de tiempo de espera de pasajeros con respecto al lazo abierto: se ve la mejora o empeoramiento de la espera por pasajero hasta antes de abordar el bus.
- Referencia *headway* promedio: se busca analizar el comportamiento de la regulación y cómo los cambios de parámetros pueden influenciar en la diferencia que se tiene según lo esperado.
- Tiempo total de *holding*: se analiza este parámetro para tener en cuenta las consecuencias no deseables de mejorar la regulación y el tiempo de espera. Se considera el tiempo total de *holding* aplicado a los pasajeros dentro del intervalo de análisis.
- Tiempo total *skipping*: se analiza este parámetro para tener en cuenta las consecuencias no deseables de mejorar la regulación y el tiempo de espera. Se considera el tiempo total de *skipping* aplicado a los pasajeros dentro del intervalo de análisis.
- Valor función objetivo: Todas las componentes anteriores se ponderan y se agrupan en un solo indicador (ecuaciones 9, 10, 11, 12 y 13), que es la función objetivo que se busca minimizar.

Entonces considerando los anteriores indicadores se realizan los siguientes análisis de sensibilidad:

1. Estudio de tamaño de selección de paraderos.
2. Análisis de pasos futuros equivalentes.
3. Análisis de *headway* de referencia.
4. Análisis de θ_1 peso de tiempo de espera, θ_2 *headway*, θ_3 *holding* y θ_4 *skipping* en función objetivo.

Estos análisis se realizan sin predictor de demanda, para obtener los mejores resultados a los cuales pueden llegar los controladores.

Los primeros tres análisis están enfocados a observar el comportamiento de los controladores bajo cambios en el diseño del controlador en sí. Finalmente el cuarto análisis se enfoca a los pesos θ dentro de la función objetivo.

3.9.1. Análisis de tamaño de selección de paraderos P^*

Este análisis se enfoca a observar el impacto de un cambio en el tamaño de la vecindad de paraderos P^* dentro del método de solución del controlador predictivo híbrido basado en selección de paraderos.

Se fijan los pesos y variables de ajuste de la función objetivo de la siguiente forma:

- θ_1 : 1 (Peso tiempo de espera de pasajeros, asociado a ecuación (9))
- θ_2 : 10 (Peso *headway*, asociado a ecuación (10))
- θ_3 : 1 (Peso *holding*, asociado a ecuación (11))
- θ_4 : 10 (Peso *skipping*, asociado a ecuación (12))
- N_p : 3 (horizonte de predicción dentro de vecindad P^*)
- \bar{H} : 4.5 minutos (intervalo de diseño que satisface la demanda de la red durante cierto periodo)

La cantidad de paraderos se varía dentro del intervalo [1 : 1 : 19]. En la Figura 14 se presentan los indicadores para analizar el comportamiento de las estrategias T_parad y S_parad . Al observar el indicador asociado al tiempo de espera (Figura 14.a), en todos los casos se obtiene un mejor resultado con S_parad , además si se aumenta el tamaño de la selección de paraderos disminuye la mejora en este indicador. En el caso del indicador asociado al error de regulación (Figura 14.b), se tienden a obtener mejores resultados con una pequeña cantidad de paraderos, además en casi todas las situaciones se obtiene un mejor resultado con S_parad que con T_parad . En relación a los indicadores de las acciones de control (Figura 14.c y 14.d), los

tiempos de *holding* y *skipping* tienden a disminuir con el aumento de la cantidad del tamaño de la selección de paraderos. Finalmente al observar el valor de la función objetivo (Figura 14.d), no se obtiene una tendencia clara de S_parad .

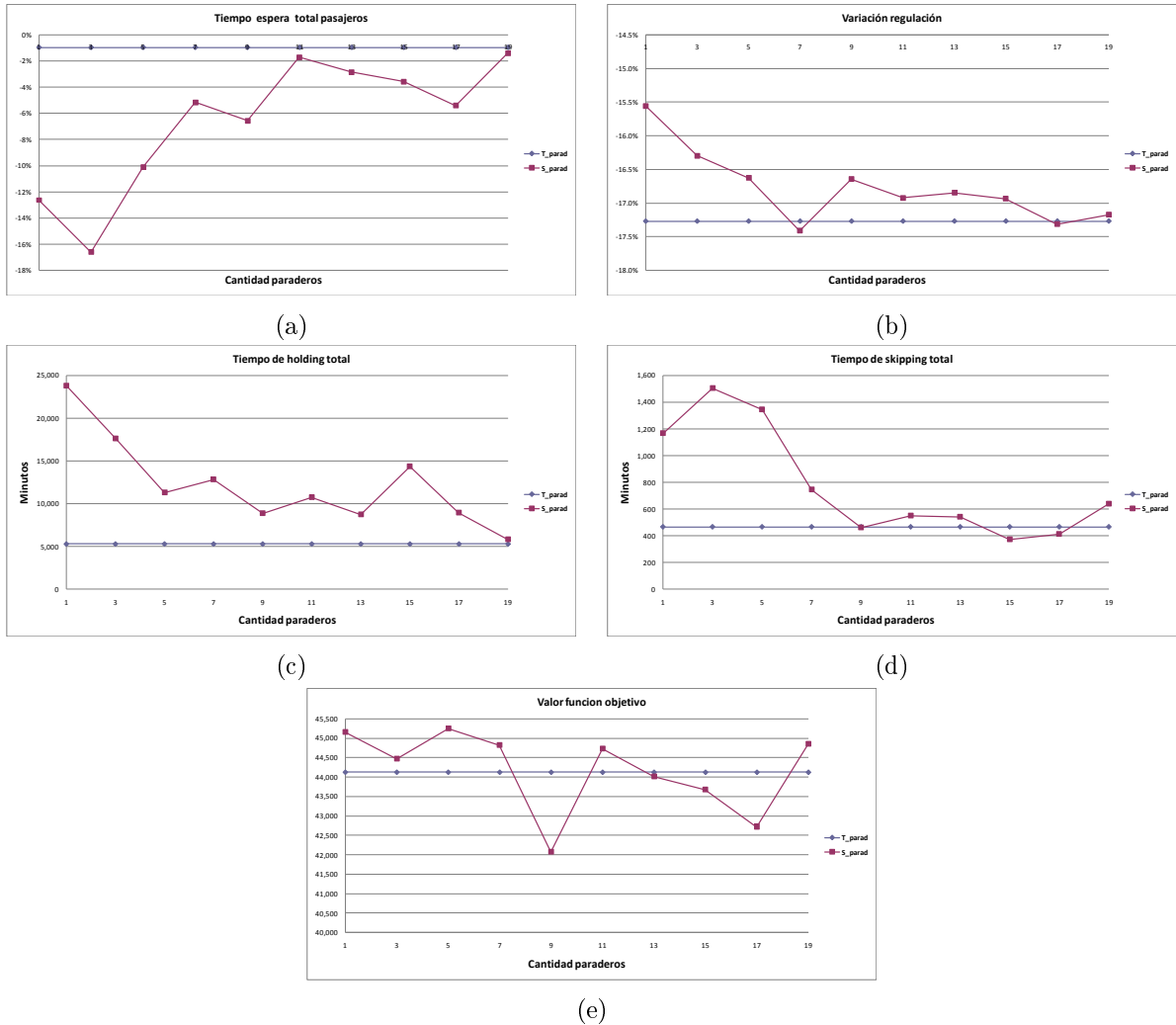


Figura 14: Variación tiempo de espera (a), Regulación promedio (b), Tiempo *holding* total (c), Tiempo de *skipping* total (d) y Valor de la función objetivo (e).

De este análisis se concluye que solamente modificar la cantidad de paraderos no entrega la información total de cómo se comporta el nuevo algoritmo S_parad en comparación a aumentar o disminuir la cantidad de paraderos, eso si se observa que el caso S_parad tiende a comportarse como el algoritmo T_parad a medida que se incrementa el tamaño de la vecindad de paraderos P^* .

3.9.2. Análisis de pasos futuros equivalentes

Se propone sensibilizar el par de variables $[P^*, N_P]$ (*vecindad de paraderos, horizonte de predicción dentro de vecindad*) para visualizar cómo se comportan los algoritmos frente a un aumento de pasos futuros de predicción; para el caso de S_parad , se calcula cuántos son los pasos equivalentes que se tendrían que predecir en el sistema completo y se utiliza como referencia para compararlo con T_parad .

Se fijan los otros parámetros de la siguiente forma:

- θ_1 : 1 (Peso tiempo de espera, asociado a ecuación (9))
- θ_2 : 10 (Peso *headway*, asociado a ecuación (10))
- θ_3 : 0.1 (Peso *holding*, asociado a ecuación (11))
- θ_4 : 5 (Peso *skipping*, asociado a ecuación (12))
- \bar{H} : 4.5 minutos (intervalo de diseño que satisface la demanda de la red durante cierto periodo)

La vecindad de paraderos P^* está en el conjunto $[1, 3, 5, 7, 9, 17]$ y los pasos futuros N_P dentro de la vecindad varían entre 2 y 5. El algoritmo T_parad se aplica directamente para $N_P = [2, 3, 4, 5]$ y coincide con los pasos futuros equivalentes del controlador. Los resultados se presentan en la Figura 15.

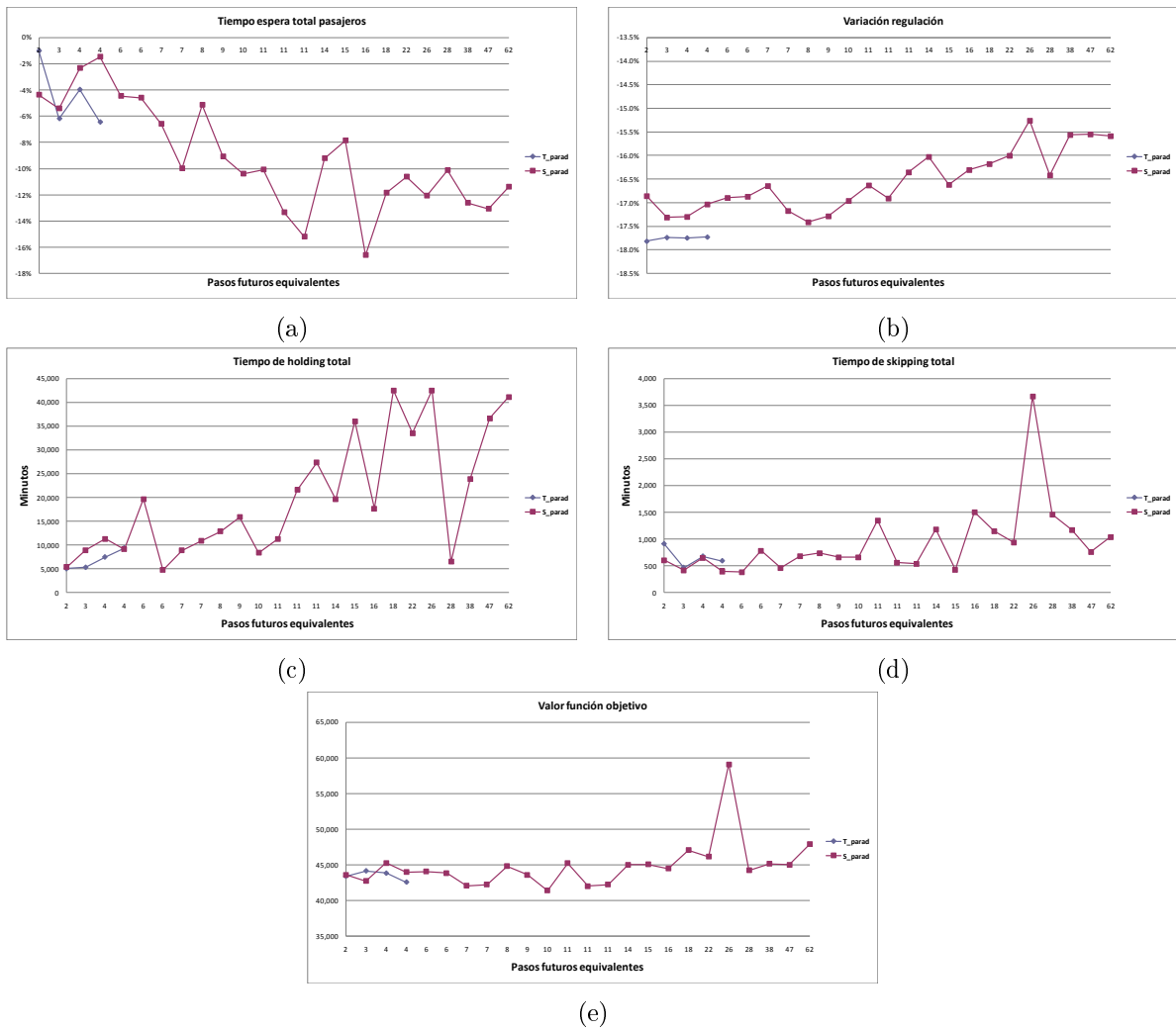


Figura 15: Variación tiempo de espera (a), Regulación promedio (b), Tiempo *holding* total (c), Tiempo de *skipping* total (d) y Valor de la función objetivo (e).

Al observar el indicador de tiempo de espera (Figura (15.a)) este tiende a disminuir con aumento de N_P , con el algoritmo S_parad se logran mejoras mucho más significativas que con T_parad . En el caso del indicador de error de regulación (Figura (15.b)), tiende a mejorar con el aumento de N_P en S_parad , en el caso de T_parad se mantiene constante. En relación a las acciones de control, El tiempo de *holding* (Figura (15.c)) de ambos algoritmos tiende a aumentar con el aumento de pasos hacia futuro, eso si, S_parad de forma oscilante; el tiempo de *skipping* (Figura (15.d)) también tiende a aumentar para ambos algoritmos pero no tan marcadamente y su oscilación es en torno a una referencia.

Finalmente al observar el valor de la función objetivo (Figura (15.e)), para el caso T_parad disminuye con el aumento de pasos hacia futuro, en cambio S_parad tiende a aumentar

levemente y de forma oscilante.

Este análisis valida el uso de el algoritmo de control S_parad porque mantiene dentro de un mismo rango la función objetivo pero tiende mejorar en el ámbito de los tiempos de espera y regulación.

3.9.3. Análisis de *headway* de referencia

A continuación se varía el *headway* de referencia \overline{H} , para observar cómo la componente de la función objetivo asociada a la regulación del *headway*(ecuación ((10))) impacta en los resultados.

Se fijan los otros parámetros de la siguiente forma:

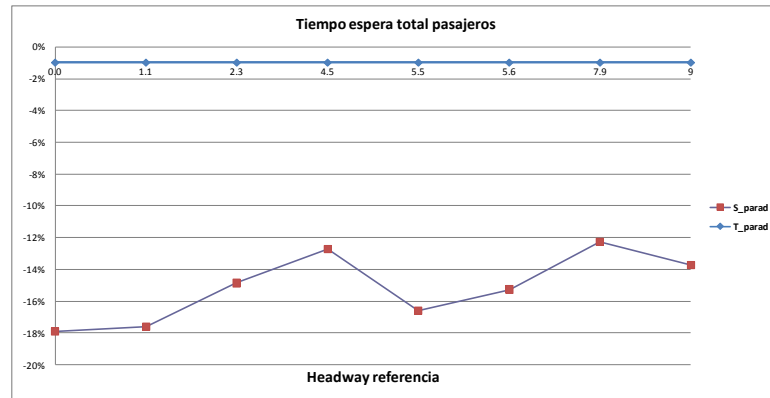
- θ_1 : 1 (Peso tiempo de espera, asociado a ecuación (9))
- θ_2 : 10 (Peso *headway*, asociado a ecuación (10))
- θ_3 :0.1 (Peso *holding*, asociado a ecuación (11))
- θ_4 : 5 (Peso *skipping*, asociado a ecuación (12))
- N_p :3 (horizonte de predicción dentro de vecindad P^*)
- P^* : 3 (vecindad de selección de paraderos)

El *headway* de referencia de diseño se definió en 4.5 [*min*] y sobre ese valor se hacen cambios porcentuales (Tabla 7) para analizar su impacto en el sistema.

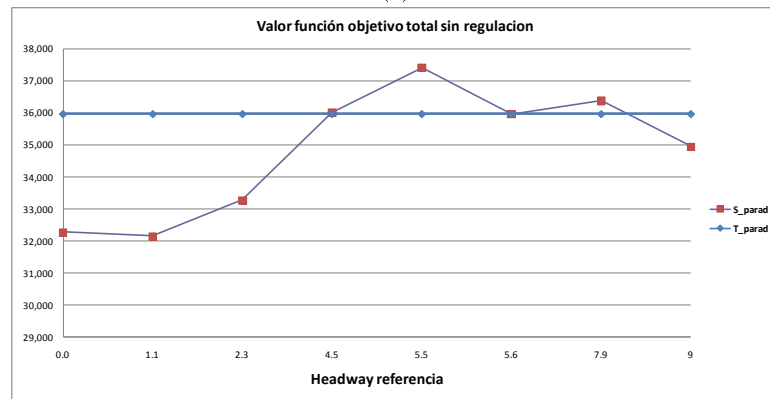
porcentaje	<i>Headway</i>
200 %	9.01
175 %	7.88
125 %	5.63
121 %	5.45
100 %	4.5
50 %	2.25
25 %	1.13
0 %	0

Tabla 7: *Headways* de referencia analizados

En la Figura (16) se presentan los resultados a analizar.



(a)



(b)

Figura 16: Variación tiempo de espera (a) y valor función objetivo (b)

El valor de la función objetivo se presenta sin la componente de la regulación ya que su valor se modifica según la variable analizada, por lo que si se incorpora los resultados no son comparables.

Para T_parad no se observa un impacto en el comportamiento tanto en tiempo de espera como en el valor de la función objetivo, en cambio para S_parad se observa una tendencia a tener mejores resultados disminuyendo el valor de \bar{H} .

3.9.4. Análisis de headway de referencia según la flota óptima

De la conclusión del análisis anterior se define realizar la misma sensibilización de headway, pero con flotas distintas ya que eso permite observar si el comportamiento se mantiene a pesar de haber cambiado el headway del sistema. Se aplica para una flota de 10 y 12

buses, considerando que el análisis previo fue sobre 11 buses. Se mantienen todos los otros parámetros.

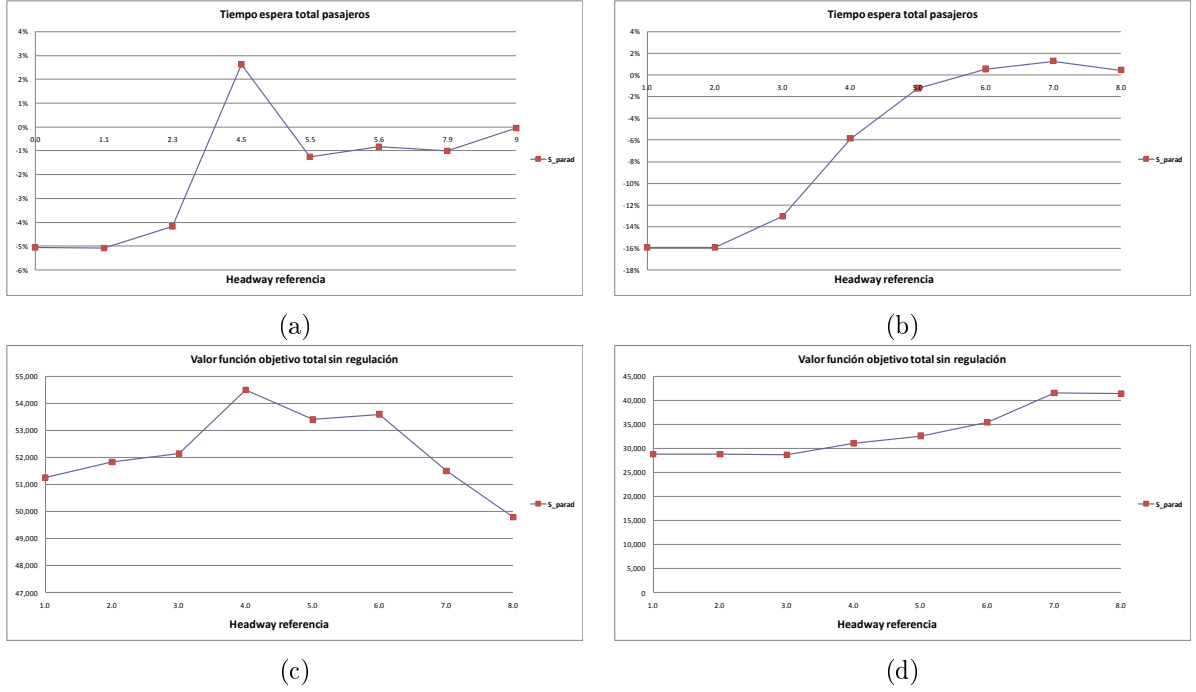


Figura 17: Variación tiempo de espera (a) y valor de función objetivo (c) para flota óptima 10; y Variación tiempo de espera (b) y valor de función objetivo (d) para flota óptima 12.

Como se puede observar en la Figura ((17)), mientras menor es el *headway* de referencia \bar{H} , menor es el tiempo de espera y menor es el valor de la función objetivo, lo cual confirma la conclusión del análisis anterior de que tener una referencia pequeña significa un mejor comportamiento en el controlador.

En relación a la función objetivo, si se fija $\bar{H} = 0$, la componente $J_2 = \sum_{l=1}^{Np} [\theta_2 \cdot (\hat{H}_i(k+l) - \bar{H})^2] |_{i=i(k+l-1)}$ (ecuación 10), queda dependiendo solamente de \hat{H}^2 y no de un polinomio, por ende, esta simplificación no cambia el impacto de las variables en si, sino que se despejan, dejando los pesos θ ponderando solamente una potencia de \hat{H} a la vez.

3.9.5. Análisis de variación de pesos θ en función objetivo

A continuación se presenta el análisis de los diferentes pesos de la función objetivo ($\theta_1, \theta_2, \theta_3, \theta_4$) presentados en la sección 3.5.

3.9.5.1 Análisis peso del tiempo de espera

Se analiza el impacto del peso de la componente del tiempo de espera en la función objetivo θ_1 . Se fijan los otros parámetros de la siguiente forma:

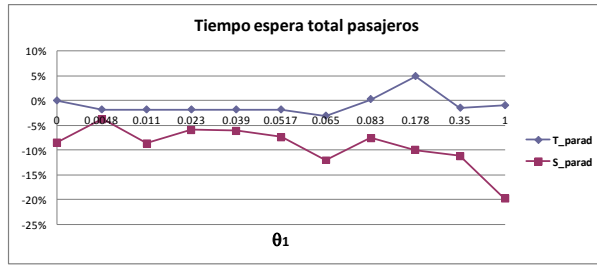
- θ_2 : 10 (Peso *headway*, asociado a ecuación (10))
- θ_3 : 0.1 (Peso *holding*, asociado a ecuación (11))
- θ_4 : 5 (Peso *skipping*, asociado a ecuación (12))
- N_p : 3 (horizonte de predicción dentro de vecindad P^*)
- P^* : 3 (vecindad de selección de paraderos)
- \bar{H} : 0 minutos (intervalo de diseño que satisface la demanda de la red durante cierto periodo)

Para el peso de tiempo de espera se toman los valores de la Tabla 8. Estos pesos buscan modificar el impacto relativo de la componente de tiempo de espera en la función objetivo.

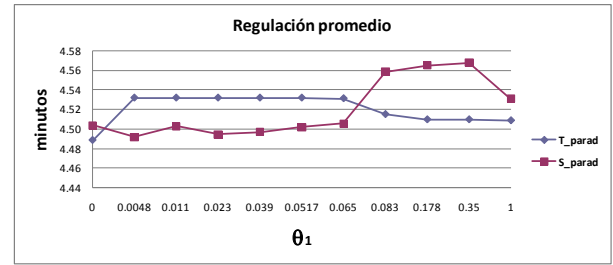
Peso tiempo espera	Peso tiempo espera	Peso tiempo espera
0	0,039	0,178
0,0048	0,0517	0,35
0,011	0,065	1
0,023	0,083	

Tabla 8: Pesos de análisis de tiempo de espera

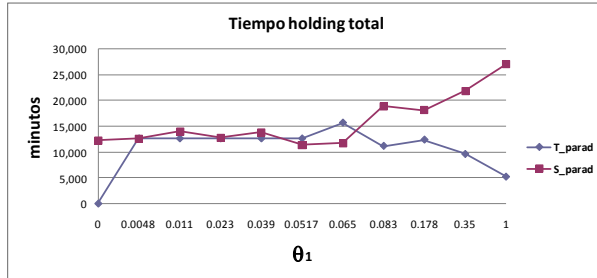
Los resultados se presentan en la Figura 18.



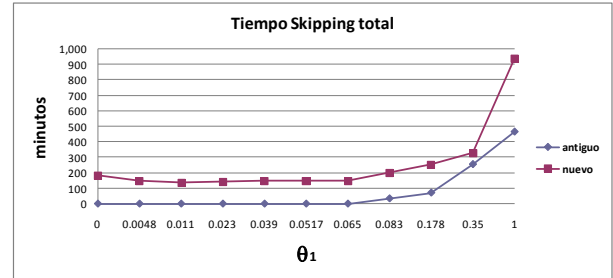
(a)



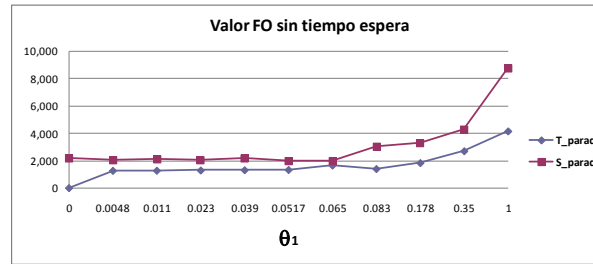
(b)



(c)



(d)



(e)

Figura 18: Variación tiempo de espera (a), Regulación promedio(b), Tiempo *holding* total(c), Tiempo de *skipping* total(d) y Valor de la función objetivo (e).

Al observar el tiempo de espera de pasajeros, para T_parad se mantiene más o menos constante, en cambio S_parad tiende a disminuir suavemente. En el caso del valor de regulación para ambos controladores, a pesar de observar tendencias visibles son de rango pequeño (inferior a 0.01 min.) por lo que se puede indicar que se mantiene sin cambios significativos. En el caso de las acciones de control, el tiempo de *holding* total se mantiene más o menos constante hasta 0,065 desde donde aumenta en forma progresiva para S_parad y comienza a disminuir para T_parad . El tiempo de *skipping* total se tiende a mantener bien constante hasta 0,065 y de ahí empieza a crecer progresivamente para ambos controladores. Finalmente el valor de la función objetivo sin tiempo de espera se mantiene constante hasta 0.065 y de ahí empieza a crecer progresivamente para ambos controladores.

Como es de esperar un aumento del peso del tiempo de espera tiende a generar menores valores de tiempo de espera, pero como consecuencia se tiene un aumento de la componente de *skipping* y que finalmente radica en que la función objetivo, para pesos muy fuertes aplicados al tiempo de espera, tiende a empeorar.

3.9.5.2 Análisis de peso de *headway*

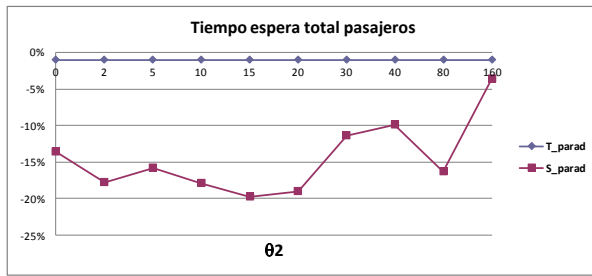
Se analiza el impacto del peso de la componente de *headway* (θ_2) en la función objetivo.

Se fijan los otros parámetros de la siguiente forma:

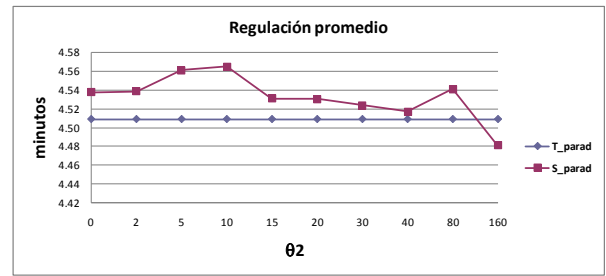
- θ_1 : 1 (Peso tiempo de espera, asociado a ecuación (9))
- θ_3 : 0.1 (Peso *holding*, asociado a ecuación (11))
- θ_4 : 5 (Peso *skipping*, asociado a ecuación (12))
- N_p : 3 (horizonte de predicción dentro de vecindad P^*)
- P^* : 3 (vecindad de selección de paraderos)
- \bar{H} : 0 minutos (intervalo de diseño que satisface la demanda de la red durante cierto periodo)

Para el peso de *headway* θ_2 se toman los siguientes valores: [0, 2, 5, 10, 15, 20, 30, 40, 80, 160].

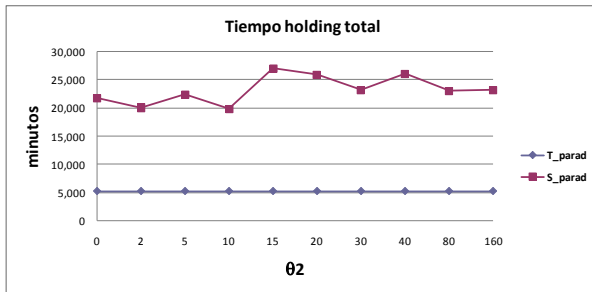
Los resultados se presentan en la figura 19.



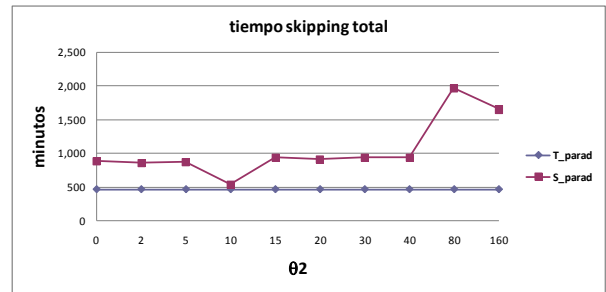
(a)



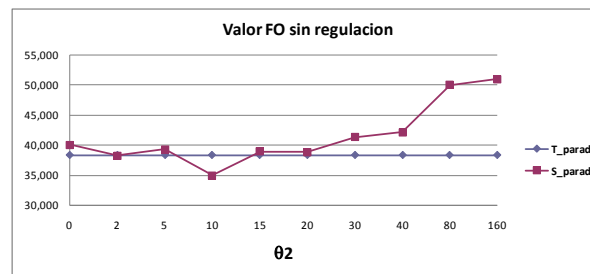
(b)



(c)



(d)



(e)

Figura 19: Variación tiempo de espera (a), Regulación promedio(b), Tiempo *holding* total(c), Tiempo de *skipping* total(d) y Valor de la función objetivo (e).

Lo primero que se observa es la insensibilidad de T_{parad} frente a cambios del peso de *headway* θ_2 , esto se debe a la forma en que se define la función objetivo original, ya que con pocos pasos no se logra interiorizar un cambio en la componente *headway*. Se requerirían varios pasos hacia futuro para que más de un bus llegase a un paradero y con eso se lograse \hat{H} .

En relación a S_{parad} , el tiempo de espera total de pasajeros se comporta más o menos constante con un valor de θ_2 inferior que 20, de ahí sube considerablemente.

Sobre el indicador de regulación, primero se debe aclarar que desde ahora no se presenta el porcentaje de variación con respecto a la referencia, sino que se presenta el valor absoluto ya que el *headway* de referencia es el valor cero. En este análisis particular la regulación tiene

una tendencia a disminuir con el aumento del peso de *headway*, lo cual es esperable. En el caso de las acciones de control, la componente *holding* tiene dos niveles de referencia, hasta un valor de 10 alrededor de 20.000, de ahí sube alrededor de 25.000. La componente *skipping* se tiende a mantener bien constante excepto para pesos grandes de *headway* donde aumenta considerablemente. Finalmente el valor de la función objetivo sin *headway* se mantiene bien constante, eso si aumenta levemente con aumento de peso desde valor 30 en adelante.

Este análisis permite concluir que realizar estrategias de control fuertemente basado en regulación de frecuencia no parece ser lo más idóneo, ya que mientras mayor es el peso que se le asigna a esta componente, peores resultados se tienen en la función objetivo y el tiempo de espera aumenta.

3.9.5.3 Análisis de peso de *holding*

Se analiza el impacto del peso de la componente de *holding* en la función objetivo θ_3 . Se fijan los otros parámetros de la siguiente forma:

- θ_1 : 0.35 (Peso tiempo de espera, asociado a ecuación (9))
- θ_2 : 10 (Peso *headway*, asociado a ecuación (10))
- θ_4 : 5 (Peso *skipping*, asociado a ecuación (12))
- N_p : 3 (horizonte de predicción dentro de vecindad P^*)
- P^* : 3 (vecindad de selección de paraderos)
- \bar{H} : 0 minutos (intervalo de diseño que satisface la demanda de la red durante cierto periodo)

Para el peso de *holding* se toman los valores de la tabla (9).

Peso <i>holding</i>	Peso <i>holding</i>	Peso <i>holding</i>
0,01	0,5	1,3
0,05	0,7	1,5
0,1	1	2

Tabla 9: Pesos de análisis de componente *holding*

Los resultados se presentan en la Figura (20).

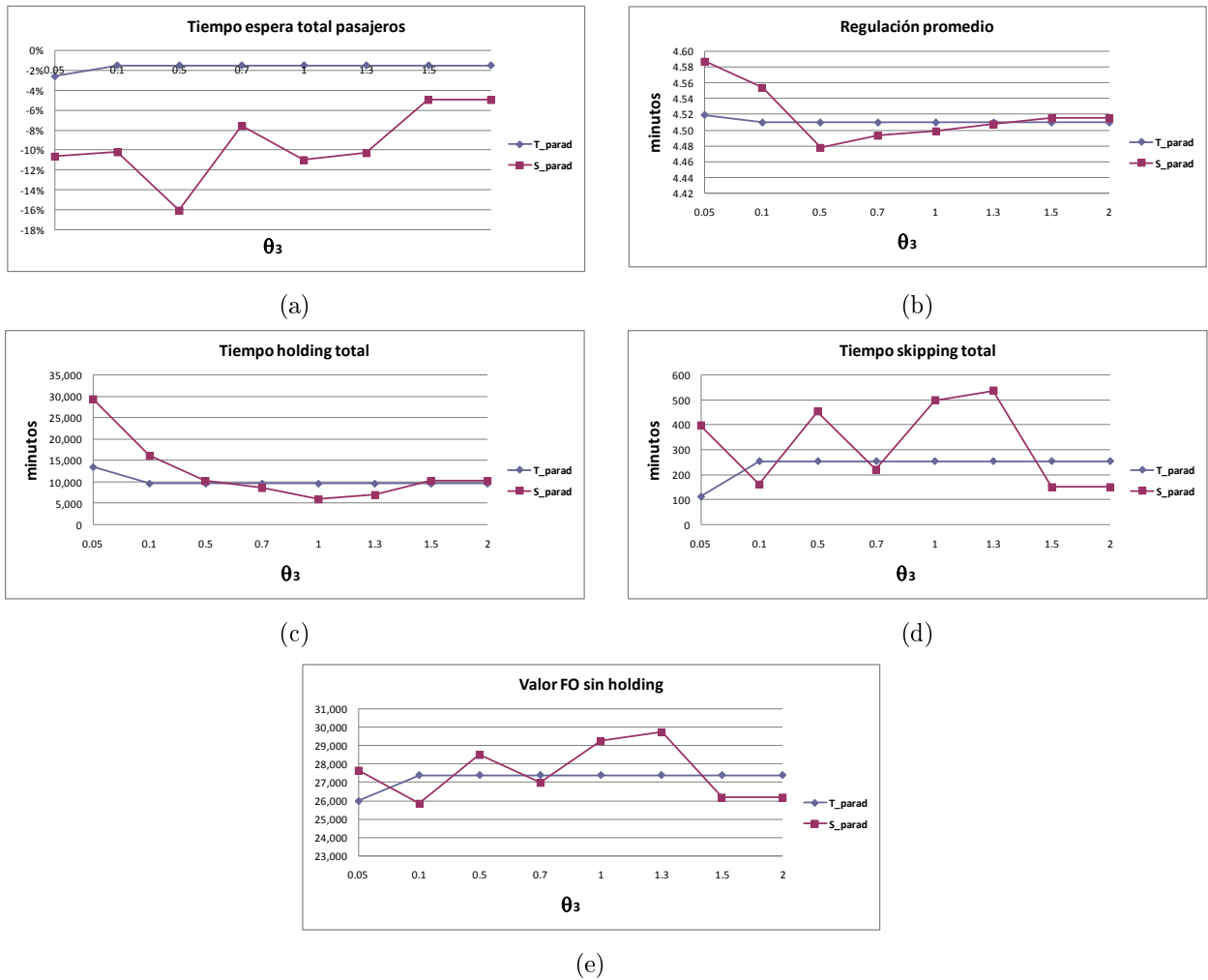


Figura 20: Variación tiempo de espera (a), Regulación promedio(b), Tiempo *holding* total(c), Tiempo de *skipping* total(d) y Valor de la función objetivo (e).

Al observar el tiempo de espera total de pasajeros, para T_parad se mantiene constante solamente con una pequeña disminución para el valor de *holding* más bajo, en cambio S_parad tiende a aumentar con el aumento del valor de θ_3 . En el caso del valor de regulación T_parad , no tiene cambios, en cambio S_parad tiende a disminuir fuertemente al inicio para después estabilizarse. En el caso de las acciones de control, la componente *holding* para ambos casos tiende a disminuir, pero de forma más marcada para S_parad , para valores mayores de θ_3 se estabiliza para ambos casos. La componente *skipping* se tiende a mantener bien constante para T_parad en cambio para S_parad se tiene un comportamiento oscilante. Finalmente el valor de la función objetivo sin componente de *holding* se mantiene constante para T_parad ,

en cambio para S_parad se tiene un comportamiento oscilante.

Para S_parad , el ir modificando el peso θ_3 en la función objetivo tiene como consecuencia que se utilice la otra acción de control para compensar estos cambios de peso relativo, por eso se obtiene un comportamiento oscilante en el caso del valor de *skipping*, ya que refleja la interrelación y complementariedad de ambas acciones de control. De la misma forma el empeorar significamente una acción de control implica que se tiene menos disponible para los ajustes, lo que redundará en un aumento del tiempo de espera.

3.9.5.4 Análisis de peso de *skipping*

Se analiza el impacto del peso de la componente de *skipping* en la función objetivo θ_4 . Se fijan los otros parámetros de la siguiente forma:

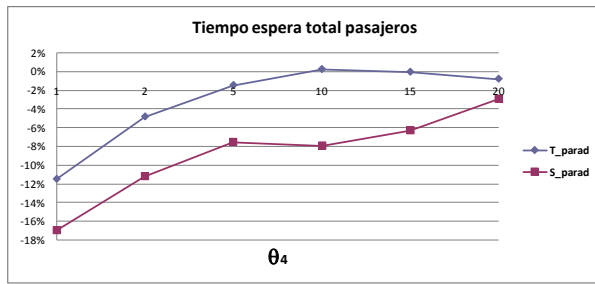
- θ_1 : 0.35 (Peso tiempo de espera, asociado a ecuación (9))
- θ_2 : 10 (Peso *headway*, asociado a ecuación (10))
- θ_3 : 0.7 (Peso *holding*, asociado a ecuación (11))
- N_p : 3 (horizonte de predicción dentro de vecindad P^*)
- P^* : 3 (vecindad de selección de paraderos)
- \bar{H} : 0 minutos (intervalo de diseño que satisface la demanda de la red durante cierto periodo)

Para el peso de *skipping* se toman los valores de la tabla (10).

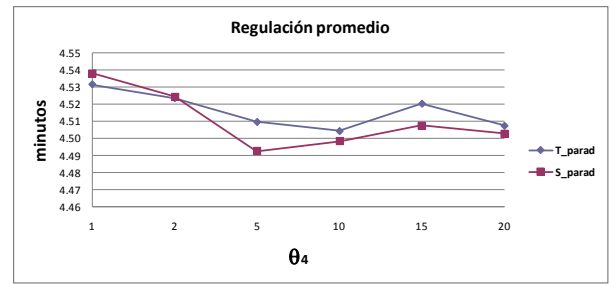
Peso tiempo espera	Peso tiempo espera	Peso tiempo espera
1	5	15
2	10	20

Tabla 10: Pesos de análisis de componente *skipping*

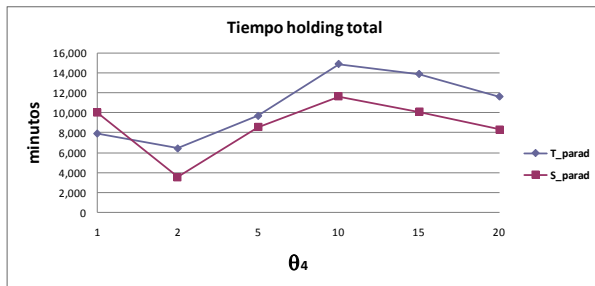
Los resultados se presentan en la Figura (21).



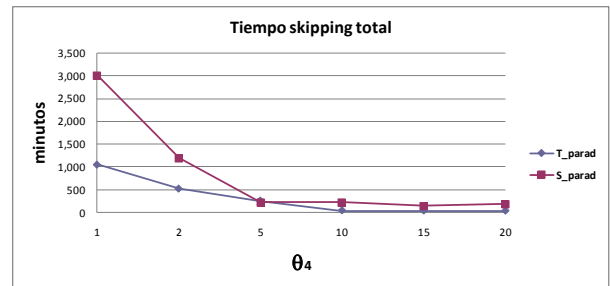
(a)



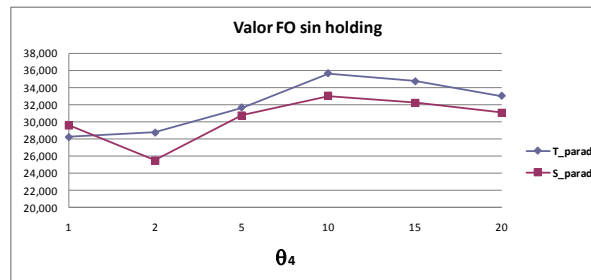
(b)



(c)



(d)



(e)

Figura 21: Variación tiempo de espera (a), Regulación promedio(b), Tiempo *holding* total(c), Tiempo de *skipping* total(d) y Valor de la función objetivo (e).

Al observar el tiempo de espera total de pasajeros, para ambos controladores se tiene un incremento si aumenta el peso θ_4 , pero con S_parad con un valor inferior siempre. En el caso de la regulación para ambos casos se tiene una disminución leve mientras aumenta el valor de θ_4 .

En el caso de las acciones de control, el tiempo de *holding* total para ambos casos presenta un aumento considerable mientras aumenta θ_4 , eso si para valores muy grandes se produce una disminución, de todas formas, siempre se tiene un menor valor para S_parad frente T_parad .

La componente de tiempo de *skipping* total obviamente tiende a disminuir para ambos casos.

Además el valor de la función objetivo sin componente de *skipping* tiende a aumentar para ambas estrategias, pero con valores muy grandes de θ_4 se tiene una disminución; S_parad en

la mayoría de los valores tiene un valor menor que T_{parad} .

Para S_{parad} , el ir modificando el peso θ_4 en la función objetivo tiene como consecuencia que se utilice la otra acción de control para compensar estos cambios de peso relativo, por eso se obtiene una curva de *holding* contraria al comportamiento de la variable *skipping*, ya que refleja la interrelación y complementariedad de ambas acciones de control. De la misma forma el empeorar significativamente una acción de control implica que se tiene menos disponible para los ajustes, lo que redundará en un aumento del tiempo de espera y un peor valor de función objetivo.

Observando el análisis de θ_3 y θ_4 se puede observar que es más simple para el sistema ajustar pesos no adecuados de las acciones de control a través *holding* que de *skipping*, lo que se puede entender dado que *holding* se comporta y se diseña de manera más flexible que el *skipping*, además sus consecuencias son más localizadas y de impacto no tan potente como lo es el *skipping*.

3.10. Discusión

En este capítulo se presenta un nuevo método de solución del controlador predictivo híbrido basado en estrategia de selección de paraderos, que permite trabajar con redes de transporte de gran tamaño, sin incurrir en un crecimiento exponencial de los tiempos de ejecución. Esto se logra mediante un cambio de enfoque que busca aplicar el control en forma local. el impacto de esta nueva metodología de solución permite disminuir los tiempos de espera y el tiempo de regulación.

Se realiza un análisis de los pesos de ajuste del controlador, concluyendo que la componente de *headway* de referencia \bar{H} en la función objetivo debe ser mínima, pero eso no implica que el tiempo de *headway* $\hat{H}(k+l)$ no sea considerado, sino que aparece en otras componentes de la función objetivo.

Finalmente se debe analizar cómo este algoritmo de control se comporta frente a cambios en la demanda, ya que una estrategia que no es lo suficientemente flexible no permite su implementación en sistemas reales, por eso se evalúa la robustez del algoritmo de control en

el capítulo 4.

4. Análisis de robustez de la estrategia de control en tiempo real con respecto a la demanda

Cuando se dispone controlar una red de transporte, se debe considerar que el comportamiento de la demanda es variable en el tiempo y no existe un modelo único para representarla, entonces el controlador debe ser flexible para poder ajustarse a estos cambios con el objetivo de seguir optimizando los indicadores de funcionamiento.

En este capítulo se estudia la robustez de las estrategias de control utilizando diferentes escenarios, lo que permitirá determinar qué característica de la demanda impacta con mayor fuerza en el comportamiento de la estrategia de control.

4.1. Descripción de escenarios

La estrategia de control HPC se aplica a un corredor simulado con las siguientes características generales:

- Número de paradas: 30.
- Largo del corredor: 10 km.
- Tiempo de simulación: 120 minutos.
- Velocidad operación promedio: 26 km/h.
- Tiempos de subida y bajada de pasajeros: 2.5 y 1.5 segundos respectivamente.
- Cada escenario dispone de 30 réplicas que corresponden a 30 días distintos.

Con el fin de evaluar y comparar el modelo propuesto bajo diferentes condiciones operaciones, se prueban dos perfiles de carga: (i) la carga del bus se concentra en el centro del corredor; (ii) la carga del bus se concentra al final del corredor. En ambos escenarios se distinguen escenarios donde: (i) la capacidad de los buses se alcanza; (ii) la capacidad de los buses nunca se alcanza; y (iii) escenarios con una alta frecuencia de servicio, es decir, *headway*

pequeño; y (iv) escenarios con una frecuencia de servicio media. Entonces, se consideran ocho escenarios que se presentan en la tabla 11.

Escenario	Capacidad de bus alcanzada	Headway de diseño	Carga sistema	Características sistema
1	Si	Corto	Centro	Headway=140 segundos; capacidad bus=100 pasajeros
2	No	Corto	Centro	Headway=120 segundos; capacidad bus=100 pasajeros
3	Si	Medio	Centro	Headway=290 segundos; capacidad bus=150 pasajeros
4	No	Medio	Centro	Headway=250 segundos; capacidad bus=150 pasajeros
5	Si	Corto	Final	Headway=140 segundos; capacidad bus=100 pasajeros
6	No	Corto	Final	Headway=120 segundos; capacidad bus=100 pasajeros
7	Si	Medio	Final	Headway=290 segundos; capacidad bus=150 pasajeros
8	No	Medio	Final	Headway=250 segundos; capacidad bus=150 pasajeros

Tabla 11: Descripción escenarios

La diferencia de los escenarios en relación al perfil de carga del sistema radica en dónde se ubican los destinos más frecuentes dentro de la matriz OD, como se observa en la figura 22. El perfil de carga medio significa que el destino de los pasajeros está concentrado en los paraderos centrales de la ruta, a su vez el perfil de carga final significa que el destino de los pasajeros está concentrado al final de la ruta.

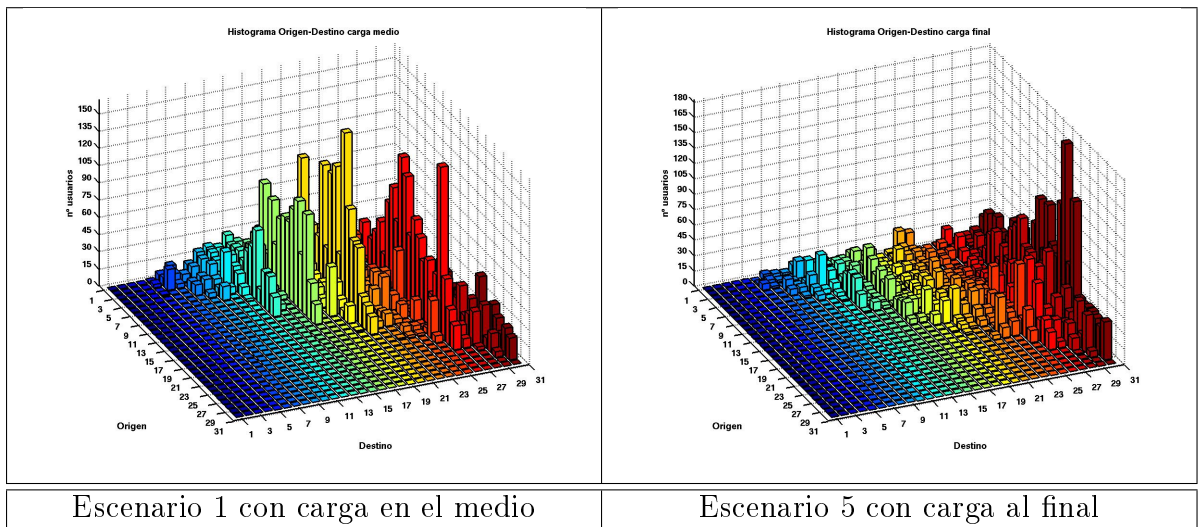


Figura 22: Comparación Matrices OD según la forma de la carga

Las diferencias asociadas a la frecuencia de servicio y si se alcanza la capacidad del bus, se reflejan en variaciones sobre el volumen de pasajeros del escenario, ya que ambas características están definidas para satisfacer el sistema. Estas diferencias se pueden apreciar en la figura 23 observando la diferencia de escala.

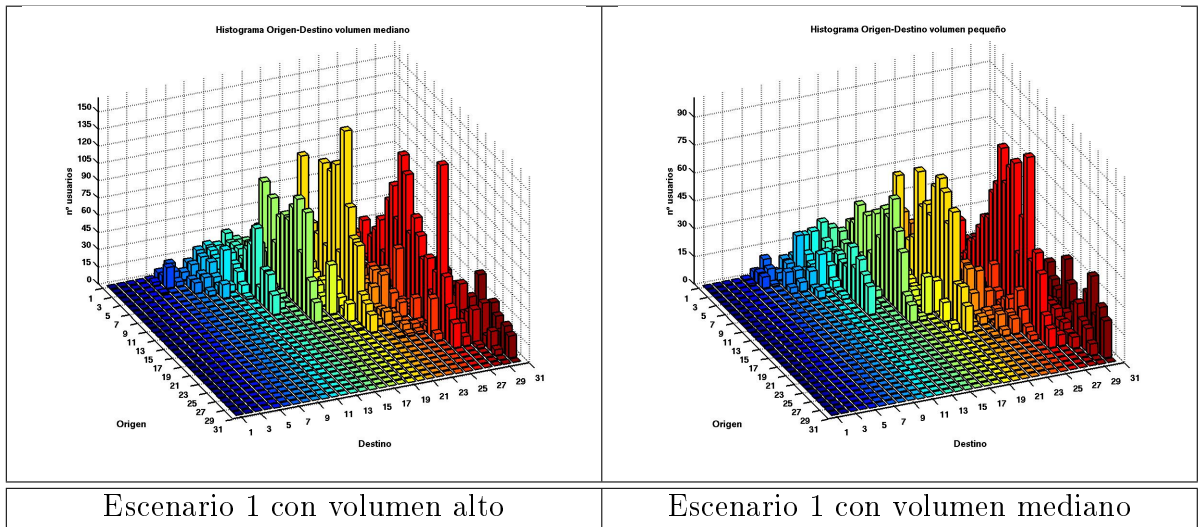


Figura 23: Comparación Matrices OD según volumen de la carga

A continuación se presentan los resultados obtenidos de la simulación.

4.2. Resultados por simulación

Se minimiza la función objetivo de la ecuación (13) usando dos estrategias. La primera estrategia es sin control y se utiliza con el propósito de comparación de caso base y la segunda estrategia de control es HPC, pero solamente utilizando la estrategia de holding.

Para cada combinación de estrategias y escenarios, se corrieron 30 réplicas que representan cada una dos horas de operación. El sistema fue simulado utilizando una adaptación del simulador desarrollado por Sáez (2007)Sáez (2007), usando números aleatorios similares y las mismas condiciones iniciales, correspondientes a buses sin carga de pasajeros iniciales y uniformemente espaciados a través del corredor. Se considera un periodo inicial de 15 minutos para todos los escenarios antes de aplicar cualquier estrategia de control.

Para cada corrida se analizan los siguientes indicadores de desempeño: tiempo de espera promedio por pasajero, incluyendo W_{first} , W_{extra} , W_{total} y R_{total} . Estos indicadores de funcionamiento se detallan a continuación:

- W_{first} : Tiempo de espera promedio experimentado por los pasajeros en una parada hasta el arribo del primer bus, prorrataada entre todos los pasajeros. Como hay un tiempo mínimo de espera dado por la frecuencia que es inevitable (es decir, la mitad del promedio de tiempo de headway), las tablas 12, 13, 14 y 15 sólo representa la diferencia entre el tiempo de espera total hasta la llegada del primer bus y el límite inferior.
- W_{extra} : Tiempo extra de espera promedio de los pasajeros que no pueden abordar el primer bus, prorrataado entre todos los pasajeros. W_{total} : W_{first} más W_{extra} .

R_{total} : W_{total} más el tiempo individual de viaje experimentado sobre el bus (incluido el tiempo de retención)

Los resultados se presentan en las tablas 12, 13, 14 y 15, tanto para el caso sin control (SC) como para el caso con Control predictivo híbrido (HPC). Las tablas incluyen el promedio y la desviación estándar de cada indicador, además del beneficio porcentual de la aplicación de la estrategia HPC sobre el caso sin control.

Escenario 1	SC	HPC	Beneficio	Escenario 2	SC	HPC	Beneficio
Se alcanza capacidad Carga al centro Frecuencia =140 seg. 100 pas/bus	min	min	%	No se alcanza capacidad Carga al centro Frecuencia =120 seg. 100 pas/bus	min	min	%
W_{first} (Dst)	0.65 (0,06)	0.24 (0,05)	62.81	W_{first} (Dst)	0.31 (0,04)	0.04 (0,02)	88.28
W_{extra} (Dst)	0,09 (0,04)	0,06 (0,03)	33.33	W_{extra} (Dst)	0,00 (0,00)	0,00 (0,00)	-
W_{total} (Dst)	0.74 (0,07)	0.30 (0,07)	59.24	W_{total} (Dst)	0.31 (0,04)	0.04 (0,02)	88.28
R_{total} (Dst)	10.01 (0,17)	9.70 (0,15)	3.10	R_{total} (Dst)	8.73 (0,10)	8.60 (0,91)	1.49

Tabla 12: Índices de desempeño para escenarios 1 y 2

Escenario 3	SC	HPC	Beneficio	Escenario 4	SC	HPC	Beneficio
Se alcanza capacidad Carga al centro Frecuencia =290 seg. 150 pas/bus			%	No se alcanza capacidad Carga al centro Frecuencia =250 seg. 150 pas/bus			%
W_{first} (Dst)	0.90 (0,40)	0.34 (0,02)	62.24	W_{first} (Dst)	0.41 (0,29)	0.10 (0,13)	75.92
W_{extra} (Dst)	0,11 (0,14)	0,11 (0,00)	0.00	W_{extra} (Dst)	0,00 (0,00)	0,00 (0,00)	-
W_{total} (Dst)	1.01 (0,54)	0.45 (0,02)	55.46	W_{total} (Dst)	0.29 (0,29)	0.10 (0,08)	75.92
R_{total} (Dst)	10.88 (0,61)	10.46 (0,10)	3.86	R_{total} (Dst)	9.21 (0,36)	9.00 (0,15)	2.28

Tabla 13: Índices de desempeño para escenarios 3 y 4

Escenario 5	SC	HPC	Beneficio	Escenario 6	SC	HPC	Beneficio
Se alcanza capacidad Carga al final Frecuencia =140 seg. 100 pas/bus			%	No se alcanza capacidad Carga al final Frecuencia =120 seg. 100 pas/bus			%
W_{first} (Dst)	0.58 (0,06)	0.14 (0,07)	75.93	W_{first} (Dst)	0.37 (0,06)	0.08 (0,05)	78.43
W_{extra} (Dst)	0,11 (0,05)	0,04 (0,02)	63.64	W_{extra} (Dst)	0,00 (0,00)	0,00 (0,00)	-
W_{total} (Dst)	0.69 (0,09)	0.18 (0,03)	73.97	W_{total} (Dst)	0.37 (0,06)	0.08 (0,04)	78.43
R_{total} (Dst)	10.48 (0,20)	10.13 (0,09)	3.34	R_{total} (Dst)	9.41 (0,16)	9.14 (0,12)	2.87

Tabla 14: Índices de desempeño para escenarios 5 y 6

Escenario 7	SC	HPC	Beneficio	Escenario 8	SC	HPC	Beneficio
Se alcanza capacidad Cargal al final Fecuencia=290 seg. 150 pas/bus			%	No se alcanza capacidad Cargal al final Fecuencia =250 seg. 150 pas/bus			%
W_{first} (Dst)	0.55 (0,26)	0.16 (0.01)	71.50	W_{first} (Dst)	0.43 (0,20)	0.12 (0.15)	72.59
W_{extra} (Dst)	0,18 (0,14)	0,11 (0.08)	38.89	W_{extra} (Dst)	0,00 (0,00)	0,00 (0,00)	-
W_{total} (Dst)	0.73 (0,39)	0.27 (0.17)	63.41	W_{total} (Dst)	0.43 (0,20)	0.12 (0.09)	72.59
R_{total} (Dst)	11.48 (0,43)	11.31 (0.29)	1.48	R_{total} (Dst)	10.11 (0,30)	9.81 (0.21)	2.97

Tabla 15: Índices de desempeño para escenarios 7 y 8

De las tablas 12, 13, 14 y 15 se presentan los análisis en relación a cómo se comporta la estrategia de control según las características de los escenarios (tabla 11).

- En todos los casos se observa una reducción de W_{first} , W_{extra} , W_{total} , R_{total} . Esto justifica el beneficio de usar estrategias control en las decisiones operacionales de una red de tránsito real.
- La estrategia de control HPC entrega mejores resultados en situaciones de frecuencia pequeña (escenarios 1, 2, 5, 6 comparado contra los escenarios 3, 4, 7, 8). Una de las razones de lo anterior, radica en que bajo esta condición ocurre una mayor cantidad de eventos en menores periodos de tiempo y espacialmente más cercanas, por lo que el controlador puede direccionar más precisamente las decisiones tomadas para minimizar la función objetivo.
- Analizando los resultados en relación al alcance de la capacidad de los buses, se tiene un mejor resultado cuando no se llega a la capacidad máxima (escenarios 1, 3, 5, 7 comparado contra los escenarios 2, 4, 6, 8). La causa está relacionada con el punto anterior, ya que si el bus se encuentra lleno, no se puede accionar sobre él en el corto plazo. Entonces si se desea también poder enfrentar esta situación, se requieren de estrategias de control que vean a más largo plazo.

- Finalmente en relación a la distribución de la carga alrededor del corredor (al final o al centro), los resultados dependen de la frecuencia de diseño. Si los escenarios poseen frecuencia pequeña, los mejores resultados se tienen con la mayor carga al final, en cambio escenarios con frecuencia alta, el mejor resultado se obtiene con la mayor carga al centro.

De estos análisis se puede concluir que la estrategia de control predictivo es lo suficientemente flexible para acomodarse al funcionamiento de diferentes escenarios, y se aprecia que tiene un mejor comportamiento cuando el headway del escenario es pequeño, los buses no llegan a capacidad máxima y la mayor demanda se tiene al final del corredor. Una forma de poder enfrentar otros tipos de escenarios, es a través de aumentar la cantidad de eventos del controlador hacia adelante que se consideran, pero como eso implica aumentar la cantidad de opciones en forma drástica, tiene que ir acompañado de una mayor capacidad de cómputo.

4.3. Discusión

En este capítulo se analiza el comportamiento de la estrategia de control presentada en el capítulo 3 frente a diferentes tipos de la demanda. Se recalca el hecho que es necesario realizar previamente una selección y ajuste de la estrategia de control seleccionada según las características del sistema de transporte que se pretende controlar, considerando longitud del corredor, cantidad de paraderos y tipo de demanda.

El método de solución del controlador predictivo híbrido basado en estrategia de selección de paraderos permite acomodarse en funcionamiento a diferentes escenarios, eso si, teniendo un mejor comportamiento para ciertos escenarios en particular, que se caractericen por tener un *headway* de diseño pequeño, cuando los buses no lleguen a su capacidad máxima y el mayor volumen de demanda se encuentre al final del corredor.

Estas demandas fueron generadas basadas en modelos de Poisson, pero se dispone de la demanda propuesta en el capítulo 2, para lo cual en el capítulo siguiente se propone un nuevo algoritmo de predicción que permita enfrentar este tipo de demanda y se utilice en la aplicación de la estrategia de control HPC.

5. Modelos de predicción de demanda origen-destino (OD)

Los modelos de predicción de demanda origen-destino en sistemas de transporte público han estado enfocados principalmente a estudios estáticos de matrices OD, porque se han utilizado para el diseño de redes de transporte, es decir, determinar el número de líneas y sus rutas asociadas. Se considera que es un estudio estático porque el deseo de movilidad de los usuarios se asume constante durante el periodo de análisis. Estas estimaciones utilizan como fuente de información conteos de tráfico promedio diario horario.

Esta tesis está enfocada a estrategias de control de sistemas de transporte público que se basan en el comportamiento dinámico de la demanda, por ende, los estudios de predicción estáticos de matrices OD no son extrapolables a los objetivos de este trabajo. Por lo anterior se revisa también la predicción de demanda asociada a gestión de tráfico, que a pesar de no enfocarse a transporte público permite proyectar sus estudios al objetivo de esta tesis. Como gestión de tráfico se entiende la optimización del sistema de transporte en su totalidad, considerando vehículos públicos y privados.

Se debe aclarar que a continuación se hace una diferencia entre *estimación* y *predicción* donde la estimación de la demanda se entiende como reconstruir su estado actual considerando un conocimiento parcial de la red, en cambio la predicción se entiende como poder conocer los estados futuros de la red considerando el estado actual y los pasados, Casas (2004).

5.1. Estimación de demanda en sistemas de transporte público

En la literatura disponible se presenta un modelo para sistemas de transporte público que se enfoca en el problema de estimación de pasajeros para matrices OD a través del conteo de pasajeros en redes congestionadas y matrices OD antiguas (Lam et al. (2002)), se extiende un enfoque de programación de dos niveles para el problema de actualización de tráfico de pasajeros OD donde el nivel superior del problema busca minimizar la suma de los errores de medición en conteo de pasajeros y matrices OD, mientras que el nivel inferior es el problema

de asignación de equilibrio estocástico del usuario para redes congestionadas. El marco de asignación de tráfico es basado en modelos de tráfico de frecuencia adaptiva, lo cual puede ayudar a determinar líneas de frecuencias de tránsito y patrones de flujo de redes que se encuentran simultáneamente congestionadas. Se adapta un algoritmo heurístico para resolver el problema de estimación de pasajeros OD.

Otro trabajo en sistemas de transporte público, Rajbhandari et al. (2003), utiliza el promedio del tiempo de pasajeros que suben y bajan de un bus y el tiempo de parada de los buses en paradero, lo cual es una información importante para estimar la capacidad del transporte público. El tiempo de parada de los buses en paradero afecta directamente los tiempos de viaje de los vehículos, y por lo tanto el tamaño de la flota necesaria para satisfacer el servicio programado. Este trabajo se centra en la estimación de tiempos de bus dwell y el impacto de subida y bajada de pasajeros en el tiempo habitado. Además se investiga el efecto de pasajeros de pie, los tiempos del día y el tipo de servicio en el tiempo de parada de los buses en paradero. Los datos se recolectaron de una base archivada, en la cual hay conteo automático de pasajeros registrada. El tiempo de parada de los buses en paradero y el promedio de tiempo de subida y bajada de pasajeros se explica utilizando estadística descriptiva.

A continuación se presentan las investigaciones existentes en la literatura asociadas a la estimación y predicción para gestión de tráfico.

5.2. Estimación y predicción de demanda en sistema de gestión de tráfico

Dentro de los trabajos revisados, Van Zuylen & Willumsen (1980) presenta un modelo inicial donde se asume que los conteos promedios de viajes en arco siguen distribución de Poisson y usan el principio de minimización de la entropía para construir un modelo de estimación de OD que permite minimizar la función de log-verosimilitud. Al mismo modelo anterior se le agrega información adicional como ser una aproximación Bayesiana en Maher (1985). Otra actualización fue agregar un estimador GLS (Generalized Least Squares) trabajo que incluye conteos de tráfico y demanda base, presentado en Cascetta (1984).

En otros modelos se intenta capturar los efectos de la congestión en la red de transporte, por medio de integrar la asignación de equilibrio en el proceso de estimación estática de demanda OD Nguyen (1977); Leblanc & Farhangian (1982); Bell (1993).

El modelo de Fisk (1989) posee un enfoque que combina el modelo de máxima entropía con un programa de asignación de equilibrio de usuario para construir un problema de programación de dos niveles .

Finalmente Florian & Chen (1993) propone una estructura de dos niveles más flexible para estimar una demanda OD consistente, donde el nivel superior es un GLS (Generalized Least Squares) basado en un modelo de estimación OD y el nivel inferior es un programa de asignación de equilibrio de usuario.

En el área de gestión de tráfico se tiene un creciente interés en la aplicación de modelos de simulación basados en asignación de tráfico dinámico (Dynamic Traffic Assignment, DTA) que ha sido acompañado por investigaciones en la estimación dinámica de matrices OD. Se considera que las condiciones de tráfico real varían con el tiempo, por eso se utiliza información en tiempo real para actualizar los modelos de predicción de demanda.

En Ziliaskopoulos & Peeta (2001) se observa que existe una revisión extensa sobre el problema de la estimación dinámica de la demanda y su inherente conexión al problema de asignación dinámica de transporte. Se investiga fuertemente el problema de estimación de demanda dinámica para poder describir la dinámica del sistema y el comportamiento del proceso en la red. Los primeros métodos propuestos estiman flujos OD dependientes del tiempo de forma aislada como el caso de intersecciones o en corredores de autopistas. Se desarrollaron cuatro métodos para identificación de flujos origen-destino dinámicos incluyendo un estimador de mínimos cuadrados, un método de optimización restringido, una fórmula simple de estimación recursiva y un método de Filtro de Kalman (Volmuller & Hamerslag (1984)). Todos estos métodos estiman, de manera separada y de forma dinámica, datos OD basados en mediciones de entrada y salida de flujos bajo el supuesto simplificado de arcos con tiempo de viaje constante. Además se proponen modelos binivel que pueden estimar el estado actual de una red y luego utilizar esta estimación como fuente para una

predicción en el mismo proceso.

Mahmassani & Zhou (2006) presentan un modelo binivel que se aplica para la operación online de asignación de tráfico dinámico en tráfico de redes a grandes escalas con diferentes fuentes de información. Esta propuesta se basa en un marco de estimación interactiva de dos niveles, donde se busca resaltar la calidad de la estimación de demanda OD por una combinación de demanda histórica estática disponible y mediciones de tráfico variable en el tiempo, en un marco de optimización multiobjetivo que minimiza la suma global de desviaciones cuadráticas. El conteo de tráfico de recorrido multi-días también se utiliza para estimar la variación de tráfico de demanda sobre múltiples días. Para evadir las dificultades de obtención de tasas de muestreo en una población de demanda, se propone una nueva estimación de demanda OD que efectivamente explota la información de distribución de demanda OD entregada por identificación automática de vehículos (Automatic Vehicle Identification, AVI) y presenta varias formulaciones robustas para considerar posibles desviaciones desde condiciones idealizadas en procesos de estimación de demanda. Se desarrolla una estimación de la estructura de demanda OD en tiempo real y un modelo de predicción con filtro de tendencia polinomial para sistemáticamente modelar patrones de demanda regular de información, desviaciones estructurales y fluctuaciones aleatorias. Así se entrega una predicción confiable que es capaz de capturar los cambios estructurales de la demanda en el tiempo variable. Se presenta además una actualización adaptiva óptima del procedimiento, basado en Filtro de Kalman, para usar estimaciones de demanda en tiempo real que permiten obtener estimaciones a priori de la media y la varianza de patrones regulares de demanda. Para mantener una representación del estado de las redes consistente con el sistema de tráfico real en un ambiente de operación en tiempo real, esta investigación propone un modelo de ajuste dinámico óptimo a la demanda OD y controladores sub-óptimos de retroalimentación eficientes para regular la demanda de entrada proveniente del simulador DTA en tiempo real, mientras se reducen las magnitudes de ajuste.

Otro trabajo binivel es presentado en Casas (2004) que propone un sistema para trabajar con la predicción de corto plazo en sistema de control en tiempo real de tráfico de transporte.

Esta propuesta está basada en redes neuronales debido a una ausencia de patrones claros en la demanda del tráfico en el corto plazo (donde es no homogénea y depende del instante de tiempo) y porque las redes neuronales permiten una estructura paralelizable lo cual satisface los requerimientos computacionales en aplicaciones de tiempo real.

Este modelo posee el inconveniente de necesitar gran cantidad de datos para el entrenamiento de la red neuronal, pero se resuelve identificando clases de rutas que conectan pares OD, luego es posible clasificar los pares OD en función de la estructura del camino que los conecta. Esto permite visualizar el problema no como la predicción de una sola gran matriz para el sistema completo, sino que se subdivide en pares de componentes donde para cada uno se diseña una red neuronal y como tiene un vector regresor menor, el entrenamiento es más fácil y factible en relación a la cantidad de datos requeridos.

Este enfoque conduce eso sí a un problema complementario: la calidad de las clases de pares OD y por lo mismo de las redes neuronales asociadas, esto se refiere a que si la clasificación no agrupa correctamente, una misma red neuronal tendrá que trabajar con patrones muy diversos por lo que no se logra disminuir la cantidad de datos de entrenamiento requeridos y la precisión es menor. La solución que se propone es obtener los recorridos por asignación dinámica de tráfico basado en micro-simulación. La asignación de tráfico heurística se convierte en la componente clave de la arquitectura propuesta porque resuelve el problema de particionar una gran red neuronal en redes neuronales más independientes y pequeñas, junto con permitir determinar el estado de la red usando microsimulación desde las matrices OD estimadas y al mismo tiempo generar los patrones en el proceso de entrenamiento de la red neuronal.

5.3. Propuesta de predicción de cantidad de pasajeros en paradero

Se presentan dos modelos de predicción los cuales buscan mejorar la calidad de la información disponible por el controlador, lo que debería mejorar la calidad de las acciones de control tomadas. El primer modelo es una modificación del predictor presentado en la sección 3.6 para tener un modelo de predicción más eficiente en el uso de recursos computacionales. La

segunda propuesta consiste en un modelo basado en redes neuronales.

5.3.1. Modelo heurístico basado en promedios para tasa de llegada constante

Este modelo se basa en el algoritmo presentado en la sección 3.6 del capítulo 3. Nace de una necesidad práctica, ya que los simuladores utilizados hasta el momento eran de rutas cortas con una cantidad acotada de paraderos (10), y generalmente 120 minutos de predicción. Como este modelo va prediciendo cada instante de llegada del siguiente pasajero, requiere de cálculo intenso que aumenta en forma proporcional a la cantidad de paraderos y tiempo de simulación, por lo que al aplicarlo al modelo de red de transporte presentado en el capítulo 2 que es basado en un corredor de 40 paraderos, el tiempo destinado a predecir la demanda lo transforma en el algoritmo no práctico. Por este motivo se realiza una modificación que está enfocada a disminuir los recursos necesarios para su cálculo sin perder su confiabilidad de resultados.

De la ecuación (18) del anexo se obtiene el siguiente instante de llegada de los pasajeros, proceso que se repite en forma recursiva hasta predecir los instantes de llegada de todos los pasajeros dentro del intervalo de análisis. Entonces la propuesta consiste en solamente predecir la llegada de un pasajero \hat{y}_{m+1} y estimar una tasa de llegada como:

$$\hat{\Delta} = \hat{y}_{m+1} - y_m = \alpha \cdot (\overline{\Delta}_{on-line}) + (1 - \alpha) \cdot (\overline{\overline{\Delta}}_{data-hist}) \quad (15)$$

Para después proyectarla como tasa de llegada constante dentro del intervalo de tiempo de análisis, esto permite evitar predecir cada instante de llegada de los pasajeros.

Esta modificación se complementa no realizando una predicción de la demanda para cada evento del sistema, sino que se define una cierta cantidad de predicciones de demanda las cuales se realizan en forma equispaciada en el intervalo de análisis.

Este algoritmo se justifica ya que el controlador no requiere de los instantes de llegada específico sino que solamente la cantidad de pasajeros que llegan entre un intervalo y otro, por ende, el proyectar la tasa promedio de un algoritmo que obtiene sus valores basado en promedios de llegadas no implica aumentar en forma fuerte el error de predicción, además la

predicción de demanda entre un evento y otro no tiene modificaciones significativas, por lo que utilizar la misma predicción de demanda para algunos eventos consecutivos no aumenta en forma considerable el error de predicción.

En el caso del destino, se mantiene el algoritmo original ya que no depende de los instantes de llegada sino de la cantidad de pasajeros previa y sus destinos asociados.

5.3.2. Modelo de predicción de cantidad de pasajeros con redes neuronales (NN)

En el capítulo 2 se presentó una nueva estrategia para generar una demanda de pasajeros que permite reflejar de forma más adecuada el comportamiento real de la demanda en un sistema de transporte público. Esta demanda se caracteriza por no poseer patrones claros de comportamiento ya que no es homogénea y depende fuertemente del instante de tiempo de análisis. Estas características dejan a las NN como candidato natural para diseñar un predictor de demanda preciso y confiable.

Cuando se utilizan las NN para la predicción de series de tiempo, generalmente se tiene un intervalo de muestreo constante y se busca saber el valor de la variable analizada en el siguiente intervalo. En este caso, se busca saber la cantidad de pasajeros que llegan al paradero dentro del siguiente intervalo definido por los instantes en que suceden los eventos en un paradero, por lo cual la duración del muestreo no es constante.

Esto entrega una dificultad adicional al tema de la predicción porque se debe definir qué variables utilizar dentro del vector regresor que permitan representar un muestreo no constante de tiempo y el valor de la variable de análisis en este intervalo. Como propuesta se define trabajar directamente con los elementos que definen el problema, es decir, instantes de tiempo de eventos y cantidad de pasajeros dentro de los intervalos definidos por los eventos. Esta propuesta es la forma más simple de definir la NN, pero genera una nueva problemática que consiste en cómo obtener información histórica para poder entrenarla ya que para obtener el instante en que suceden los eventos se requiere simular la red de transporte con controlador y predictor, por lo que no es posible contar con instantes de eventos sin disponer con antelación de un predictor.

Para superar este problema, se propone obtener instantes de eventos y sus respectivas cantidades de pasajeros de simulaciones de la red de transporte realizadas directamente con la demanda real y sin predictor, es decir, como si el controlador fuese capaz de ver cómo se comportan los pasajeros en el futuro con error cero. Esta idea es aceptable ya que el controlador tomará las mejores acciones de control y por ende los instantes de eventos serán los más ideóneos para obtener los mejores indicadores de desempeño. Entonces, si el predictor propuesto es aceptable sus resultados deben acercarse al comportamiento del controlador con demanda real.

Diseño de la estructura de la NN

Se define utilizar como estructura del predictor el modelo de red neuronal autoregresivo con entrada externa NNARX por ser una estructura estable y simple, capaz de considerar una entrada externa (instantes de eventos) y una salida de predicción en tiempo pasado (cantidad de pasajeros que llegan en los intervalos), esta estructura se prefiere cuando el sistema es determinista como es este caso. El vector regresor de la entrada se compone como sigue:

$$\varphi(k_p) = [\delta_p(k_p - 1), \dots, \delta_p(k_p - n), t(k_p), \dots, t(k_p - r)]^T \quad (16)$$

Donde:

- k_p corresponde a los eventos que suceden solamente en el paradero p .
- $\delta_p(k_p)$ corresponde a la cantidad de pasajeros que llegan entre los eventos k_p y $k_p - 1$: en la estructura de la NN se ajusta a la salida retroalimentada.
- $t(k_p)$ corresponde al instante de tiempo cuando sucede el evento k_p ; en la estructura de la NN se ajusta a la entrada externa.

La red neuronal permite obtener el valor de la cantidad de pasajeros basándose en los valores del vector regresor aplicados.

Es importante ajustar de forma adecuada el largo del vector regresor, ya que si es muy pequeño la dinámica del sistema no será modelada, pero muy grande también puede ser un

problema, según Nørgaard et al. (2000), la cantidad de pasos hacia atrás de las componentes del vector regresor (es decir, el valor de n y r) se define analizando los coeficientes de Lipschitz (ver sección 6 en Anexos).

En la figura 24 se presenta el valor del coeficiente de Lipschitz para diferentes valores de tamaño de componentes de n y r del vector regresor. Según esta tabla se escojen $n = 5, 6$ como tamaño de pasos hacia atrás de la cantidad de pasajeros y $r = 2, 3, 4$ para el tiempo de los eventos como buenos valores para representar la dinámica del sistema.

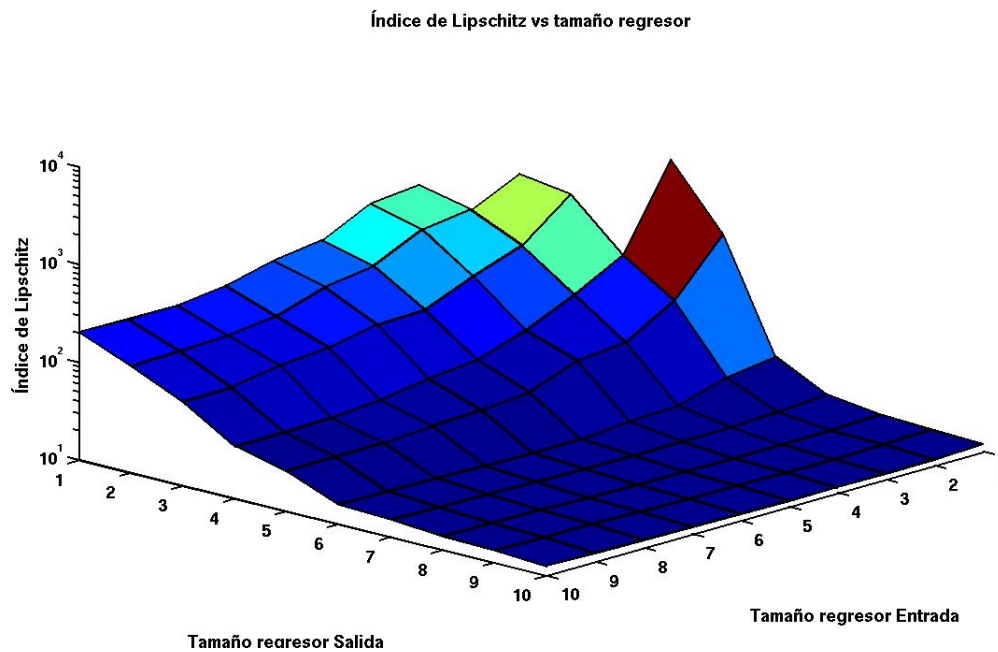


Figura 24: Coeficiente de Lipschitz para diferentes valores de regresor

Como no se tiene una sola combinación de n, r se debe ver la capacidad de predicción de todas las redes neuronales disponibles; de esta forma se definen los pares $NN[n, r]$ que representan todas las combinaciones posibles de (n, r) seleccionadas anteriormente con las cuales se construye una red neuronal por cada uno: $NN[5, 2]$, $NN[5, 3]$, $NN[5, 4]$, $NN[6, 2]$, $NN[6, 3]$, $NN[6, 4]$. Estas redes se entrenan a través del método de Levenbert-Marquardt para seleccionar la red de mejor desempeño (detalle en la sección 6 de Anexo).

Información disponible para obtención de NN

Para poder entrenar la red neuronal se disponen de 25 días los cuales se separan en dos

grupos:

- 21 días para el entrenamiento de la NN.
- 4 días para validar la predicción.

Esta división se realiza para evitar sobreajustar la red neuronal a un conjunto de datos y no disponga de la flexibilidad adecuada para otros días (al corroborar y entrenar con los mismos datos).

Los días disponibles para el entrenamiento no se consideran como un grupo de réplicas de un conjunto de datos, sino que se utilizan como una sola serie de datos que se conectan en forma consecutiva. Esto permite superar el tema de las predicciones en instantes iniciales donde la información disponible no cuenta con datos previos para el regresor.

5.3.3. Error de predicción directo

El trabajar con estrategias de control basado en HPC implica disponer de un bloque de predicción de demanda para proyectar hacia futuro cómo será el comportamiento del sistema de transporte, entonces el objetivo final del módulo de predicción es obtener mejores indicadores de desempeño de la aplicación de la estrategia de control. Bajo la explicación previa, para evaluar efectivamente si un predictor tiene un buen o mal comportamiento, se deberían analizar los indicadores de desempeño del sistema de transporte simulado con la estrategia de control basado en HPC. Realizar este proceso para cada iteración del predictor es muy costoso en tiempo, además que se necesita entrenar una red neuronal por paradero, por lo que se busca una alternativa que permita entrenar y analizar el predictor propuesto. Por esto se define analizar el rendimiento del predictor a través del error de predicción directo, esto significa evaluar cuánto difiere el estimador $\hat{\delta}_p$ con respecto al valor esperado δ_p utilizando la siguiente medida de error:

$$error_p = \frac{\sqrt{\sum_{i=1}^n (\hat{\delta}_{p,i} - \delta_{p,i})^2}}{\sum_{i=1}^n \delta_{p,i}} \quad (17)$$

La parte superior del $error_p$ entrega una estimación de la cantidad de pasajeros en que el predictor no coincidió con respecto a la cantidad correcta de pasajeros que arriban y se divide por el total correcto de pasajeros que llegan para expresar el error de forma porcentual con respecto al total. Esto último permite comparar el error entre paraderos.

El $error_p$ se aplica hasta un valor n que representa todos los datos de los 4 días que se destinan para entrenar los predictores, entonces finalmente se obtiene un vector de error \overrightarrow{error} de largo P con una componente representativa por paradero.

La comparación entre las distintas configuraciones de predictores indicadas en la sección de diseño de la estructura NN se realiza en dos fases.

Primero se comparan los diseños representándolos a través de un solo valor que es la sumatoria de las componentes del vector \overrightarrow{error} para poder realizar una primera selección de configuraciones de predictores (el resultado se presenta en la figura 25). Esta selección descarta predictores que en forma global tienen un comportamiento no óptimo para la predicción de todos los paraderos en su conjunto. De este primer análisis se seleccionan los modelos con regresores $NN[5, 2]$, $NN[5, 3]$ y $NN[6, 2]$ y además se mantiene el modelo heurístico basado en promedios, Sáez (2007), ya que representa la alternativa que se propone mejorar.

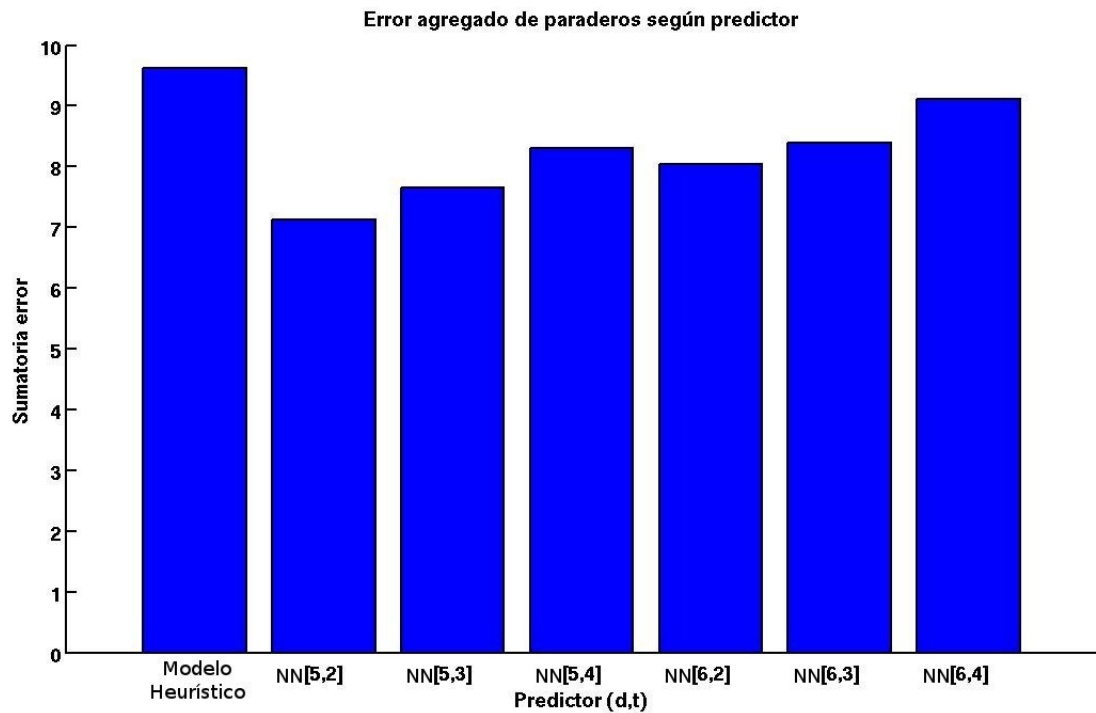


Figura 25: $Error_p$ agregado según predictor

Segundo, se revisa el vector \overrightarrow{error} para evitar tener errores significativos en algunos paraderos (el resultado se presenta en la figura 26), permitiendo escoger una configuración de predictor que tenga un buen comportamiento global y local. De este análisis se descartan los predictores NN[5,3] y NN[6,2] por tener un comportamiento promedio inferior que la alternativa NN[5,2] además de que en el paradero 20 tienen un comportamiento peor que el modelo heurístico basado en promedios.

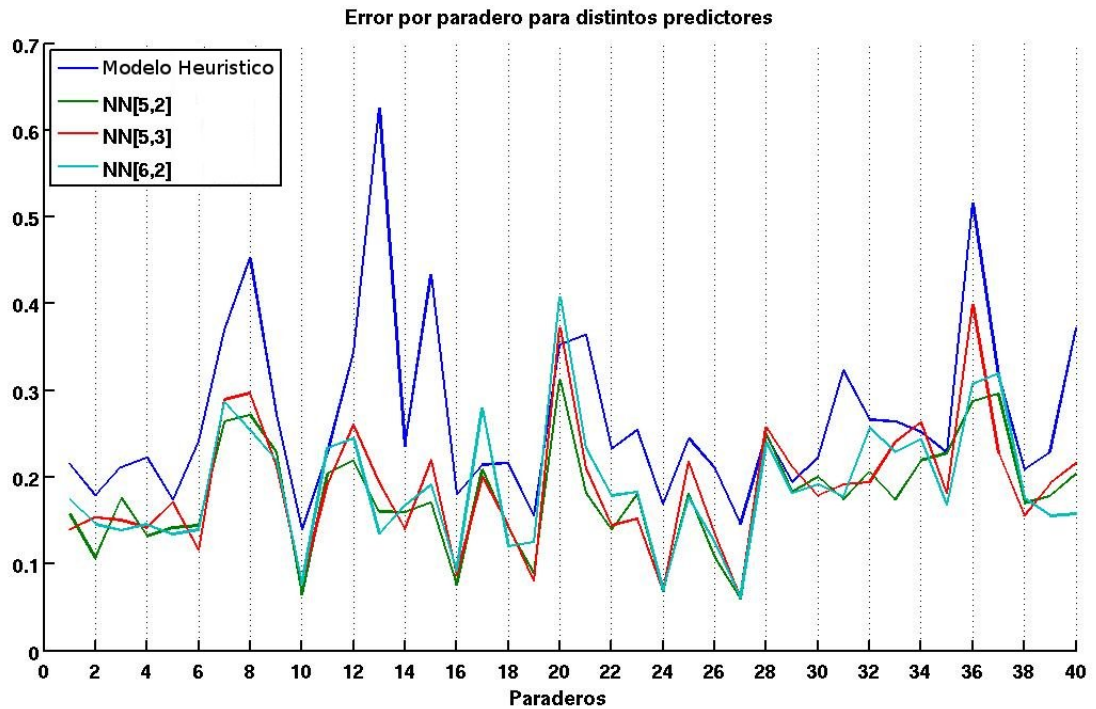


Figura 26: $Error_p$ para distintos paraderos

5.4. Análisis y sensibilización de parámetros para modelo de predicción heurístico

Para el modelo heurístico basado en promedios se debe realizar una sensibilización de los parámetros de ajuste del algoritmo, con el objetivo de mejorar la predicción del sistema. En el caso del modelo basado en NN, este ajuste de los parámetros queda realizado en la fase de entrenamiento de la red.

Los parámetros a sensibilizar son los siguientes:

1. Cantidad de predicciones realizadas: está asociado a la cantidad de veces que se debe aplicar el predictor dentro del intervalo de simulación para eventos consecutivos.
2. Cantidad de días anteriores de predictor δ_2 .
3. Valor peso α .
4. Tamaño ventana predicción d_{1a} .

El detalle y utilización en el algoritmo de predicción de estas variables se presentan en los anexos en la sección 6.

Los indicadores de desempeño analizados son los siguientes:

- Tiempo de espera total de pasajeros (Está asociado a la componente J_1 (9) de la función objetivo (13))
- Regulación promedio
- Tiempo holding total
- Tiempo skipping total

Cantidad de predicciones realizadas

En este análisis se busca ajustar la cantidad de predicciones de demanda que se realizan en el sistema para evitar realizar predicciones que tienen poco impacto basándose en que entre dos eventos consecutivos la variación de la predicción de la demanda es mínima para realizarla nuevamente.

El valor de los otros parámetros se presenta en la tabla 16.

Parámetro	Valor
θ_1 Peso tiempo de espera de pasajeros	1
θ_2 Peso headway	15
θ_3 Peso holding	0.1
θ_4 Peso skipping	5
N_p horizonte de predicción dentro de vecindad P^*	3
P^* vecindad de selección de paraderos	3
\bar{H} intervalo de diseño	0
δ_2 Cantidad de días anteriores de predictor	2
$d_{1\alpha}$ Tamaño ventana predicción	10 min.
α	0.3

Tabla 16: Valores de parámetros

Se ejecuta el algoritmo de predicción con diferentes valores de cantidad de predicciones: 1, 2, 5, 10, 15, 20, 25, 30, 50, 100, 150, 300.

La Figura 27 presenta los resultados para el método de solución del controlador predictivo híbrido basado en estrategia de todos los paraderos (T_parad), el método de solución del controlador predictivo híbrido basado en estrategia de selección de paraderos (S_parad) (presentados en el capítulo 3), y además el comportamiento de S_parad con demanda real (denominada Ref), esta última componente se considera como referencia del mejor comportamiento posible de obtener. Al observar el tiempo de espera total de pasajeros (figura 27.a) como es de esperar al utilizar el predictor los resultados se empeoran en ambos controladores, eso si S_parad tiende a tener un mejor resultado. En el caso de regulación promedio (figura 27.b), T_parad se comporta de manera similar que la referencia, en cambio S_parad tiene un valor más alto y oscila en un rango mayor.

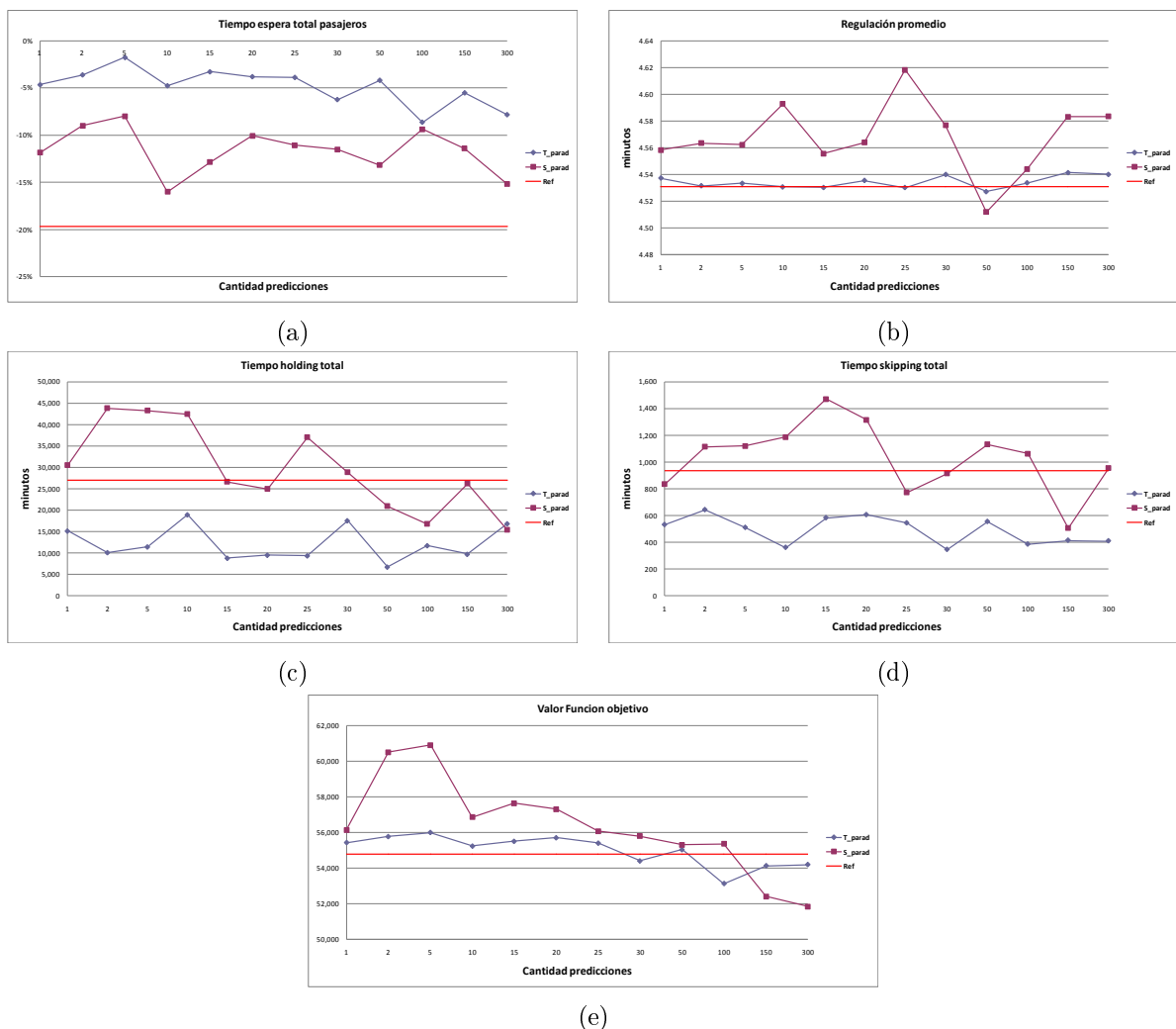


Figura 27: Tiempo de espera total pasajeros (a); Regulación promedio (b); Tiempo de holding total (c); Tiempo de skipping total (d) y Valor de la función objetivo (e).

En relación a las acciones de control, T_parad aplica una menor cantidad de tiempo de holding que la referencia, al contrario de S_parad que aplica un mayor volumen. En la componente de skipping la tendencia es similar a la componente de hoding.

A su vez al observar el valor de la función objetivo (figura 27.e), T_parad tiene un valor similar a la referencia, por lo que su rango de comportamiento es similar a disponer de demanda real, la diferencia radica en que el tiempo de espera empeora, En cambio con S_parad se tiene un valor mayor de función objetivo pero se dispone de un tiempo de espera menor.

Como análisis final, se puede indicar que el algoritmo de predicción, mientras más veces se aplique al sistema, mejores resultados se obtienen, mostrándose una tendencia a acercarse al caso con demanda real, lo cual es lo esperable. Además T_parad al tener error en el predictor, presenta un comportamiento que tiende a disminuir la cantidad de acciones de control (no se arriesga a aplicar acciones de control erróneas), lo que repercute en tener mayores tiempos de espera pero se mantiene dentro del orden el valor de la función objetivo.

Análisis de cantidad días anteriores de predictor δ_2

Se busca analizar el comportamiento del predictor al observar el impacto de modificar una de sus componentes de ajuste. Este análisis solamente se aplica en el controlador S_parad , porque es el algoritmo que obtuvo mejores resultados en la sección 3.9.

El estudio se realiza variando las tres componentes de ajuste del predictor y se promedian los resultados según la variable que se analice, en este caso, tanto la ventana de predicción d_{1a} y ponderador α tienen dos valores cada uno y se promedian los 4 resultados de aplicar un valor específico de cantidad de días anteriores δ_2 .

El valor de los otros parámetros se presenta en la tabla 17.

Parámetro	Valor
θ_1 Peso tiempo de espera de pasajeros	0.35
θ_2 Peso headway	10
θ_3 Peso holding	0.7
θ_4 Peso skipping	5
N_p horizonte de predicción dentro de vecindad P^*	3
P^* vecindad de selección de paraderos	3
\bar{H} intervalo de diseño	0
d_{1a} Tamaño ventana predicción	10-30 min.
α	0.3-0.5
Cantidad de predicciones	30

Tabla 17: Valores de parámetros

Para δ_2 se consideran los valores: 1, 2, 3, 5, 7, 9.

Los resultados de indicadores se presentan en la tabla 28. Al observar el tiempo de espera total de pasajeros (figura 28.a), este tiende a aumentar con aumento de los días. En el caso de la regulación promedio (figura 28.b) a pesar de tener una oscilación su rango es mínimo por lo que se puede considerar constante. En relación a las acciones de control, la componente tiempo holding total (figura 28.c) tiende a disminuir con el aumento de días y la componente de tiempo skipping total (figura 28.d) se comporta de manera similar, pero con una mayor oscilación. Finalmente al observar la función objetivo, se tiene una mejora con el aumento de días de estudio (figura 28.e).

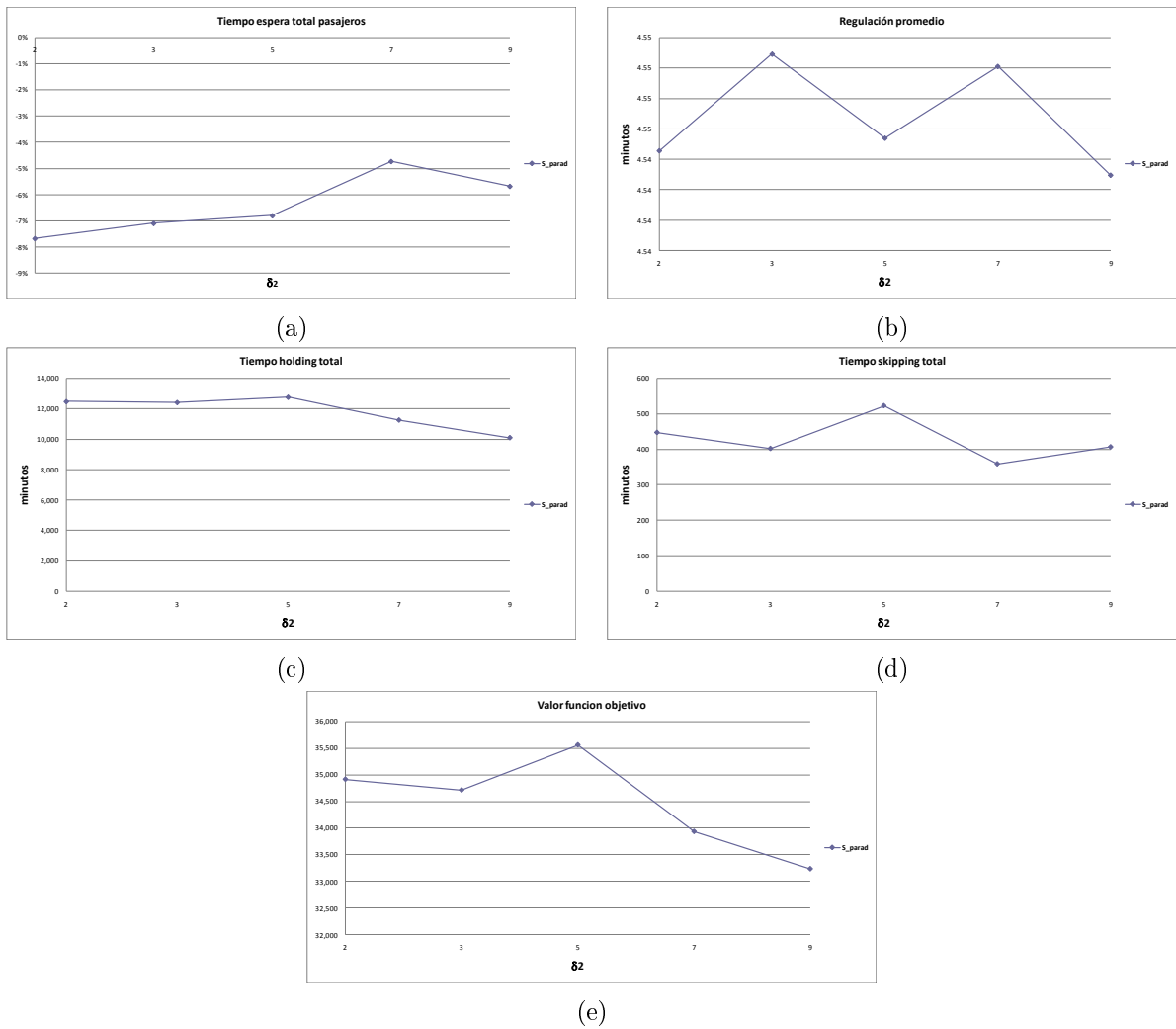


Figura 28: Tiempo de espera total pasajeros (a); Regulación promedio (b); Tiempo de holding total (c); Tiempo de skipping total (d) y Valor de la función objetivo (e).

El hecho de aumentar el valor de δ_2 implica mejorar la componente de la predicción asociada a la demanda histórica, lo cual produce un mejor desempeño global pero implica empeorar levemente el tiempo de espera total de pasajeros.

Análisis de peso predictor α

Se busca analizar el comportamiento del predictor al observar el impacto de modificar uno de sus componentes de ajuste. Este análisis solamente se aplica en el controlador S_{parad} , porque es el algoritmo que obtuvo mejores resultados en la sección 3.9.

El estudio se realiza variando las tres componentes de ajuste del predictor y se promedian

los resultados según la variable que se analice, en este caso, tanto δ_2 y d_{1a} tienen dos valores cada uno y se promedian los 4 resultados de aplicar un valor específico de cantidad de días anteriores.

El valor de los otros parámetros se presenta en la tabla 18.

Parámetro	Valor
θ_1 Peso tiempo de espera de pasajeros	0.35
θ_2 Peso headway	10
θ_3 Peso holding	0.7
θ_4 Peso skipping	5
N_p horizonte de predicción dentro de vecindad P^*	3
P^* vecindad de selección de paraderos	3
H intervalo de diseño	0
δ_2 Cantidad de días anteriores de predictor	2
d_{1a} Tamaño ventana predicción	10-30 min.
Cantidad de predicciones	30

Tabla 18: Valores de parámetros

Para los posibles valores de la componente α se considera la tabla siguiente.

Valor α	Valor α	Valor α
0	0,3	0,7
0,1	0,5	1

Tabla 19: Valor componente α

En la Figura 29, al observar el tiempo de espera total de pasajeros (figura 29.a), se mantiene constante independiente del valor de α . En el caso de la regulación promedio (figura 29.b) también se tiende a mantener constante (variación inferior a 0,03 minutos). En relación a las acciones de control, la componente tiempo holding total (figura 29.c) tiende a disminuir levemente con un mayor valor de α y la componente de tiempo skipping total (figura 29.d) se comporta de manera similar. Finalmente al observar el valor de la función objetivo (figura 29.e), se tiene una mejora con el aumento del valor de α .

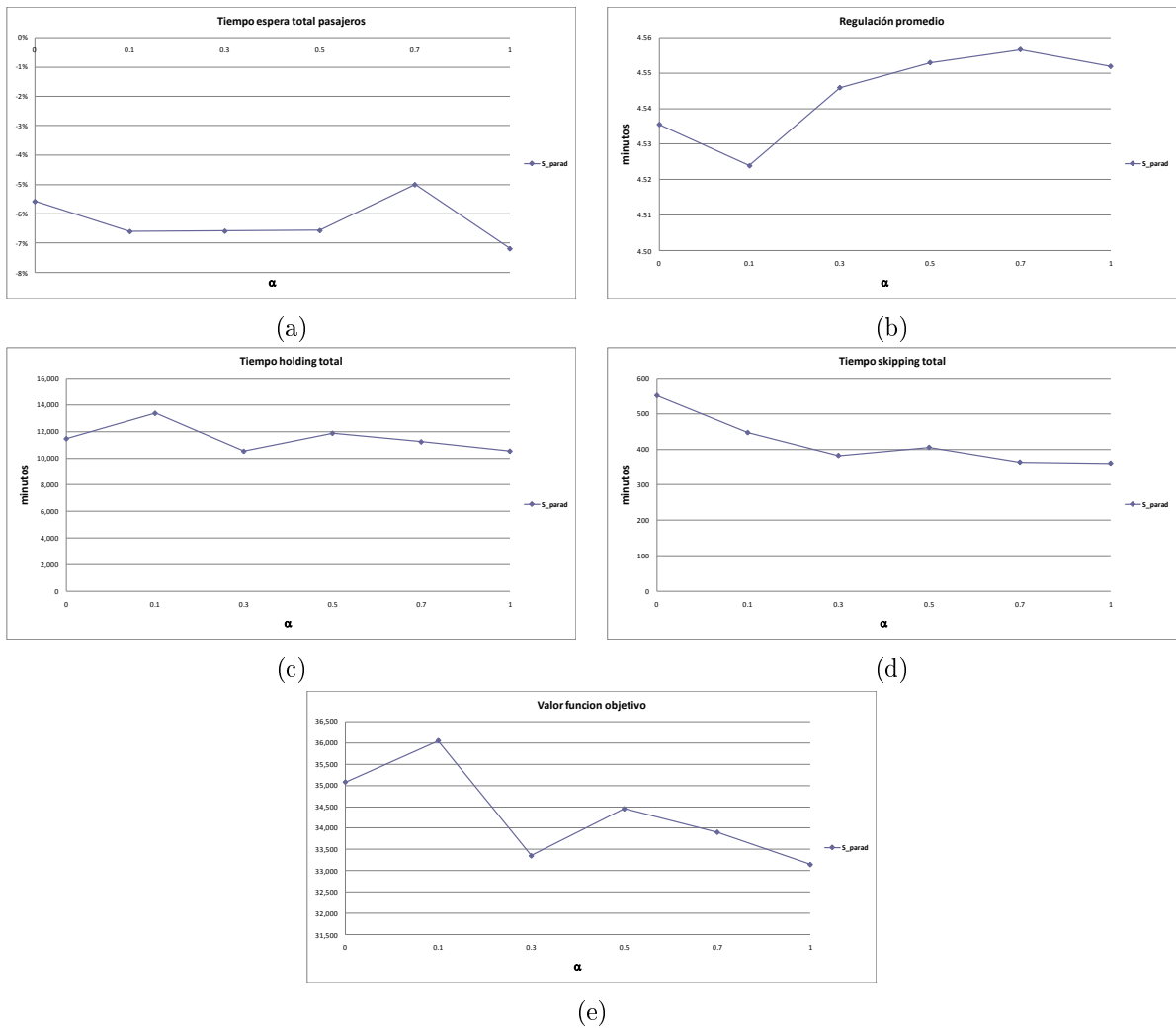


Figura 29: Tiempo de espera total pasajeros (a); Regulación promedio (b); Tiempo de holding total (c); Tiempo de skipping total (d) y Valor de la función objetivo (e).

Esto indica que considerar un mayor porcentaje de demanda on-line permite obtener mejores resultados globales en la aplicación de las acciones de control, pero que no significa que se mejore el tiempo de espera total de pasajeros.

Análisis de peso d_{1a}

Se busca analizar el comportamiento del predictor al observar el impacto de modificar una de sus componentes de ajuste. Este análisis solamente se aplica en el controlador S_parad , porque es el algoritmo que obtuvo mejores resultados en la sección 3.9.

El estudio se realiza variando las tres componentes de ajuste del predictor y se promedian los

resultados según la variable que se analice, en este caso, tanto la cantidad de días previos δ_2 y α tienen dos valores cada uno y se promedian los 4 resultados de aplicar un valor específico de tamaño de ventana.

El valor de los otros parámetros se presenta en la tabla 20.

Parámetro	Valor
θ_1 Peso tiempo de espera de pasajeros	0.35
θ_2 Peso headway	10
θ_3 Peso holding	0.7
θ_4 Peso skipping	5
N_p horizonte de predicción dentro de vecindad P^*	3
P^* vecindad de selección de paraderos	3
H intervalo de diseño	0
δ_2 Cantidad de días anteriores de predictor	3-7
α	0.3-0.5
Cantidad de predicciones	30

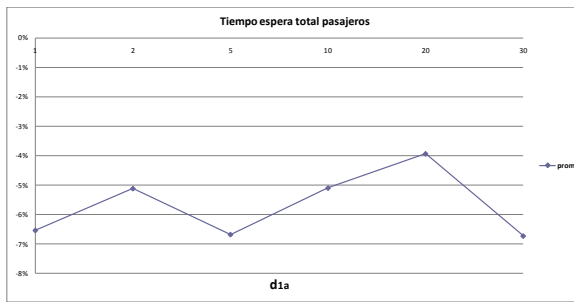
Tabla 20: Valores de parámetros

Para los posibles tamaños de ventana se considera la tabla siguiente.

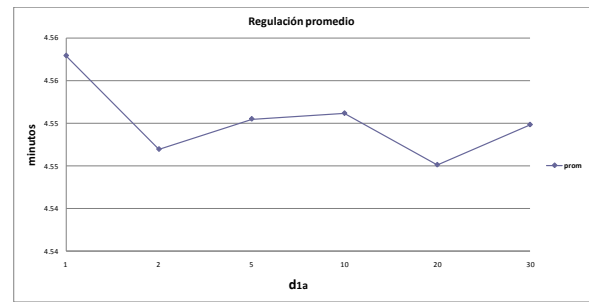
Tamaño d_{1a} (min)	Tamaño d_{1a} (min)	Tamaño d_{1a} (min)
1	5	20
2	10	30

Tabla 21: Tamaño d_{1a}

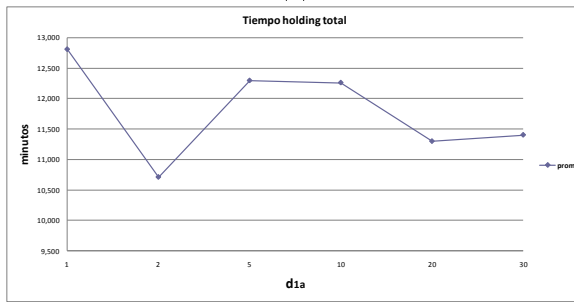
A continuación se presentan los resultado del análisis.



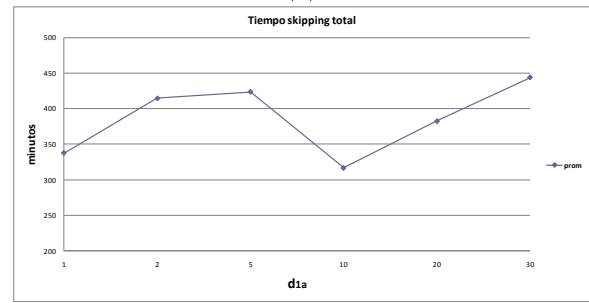
(a)



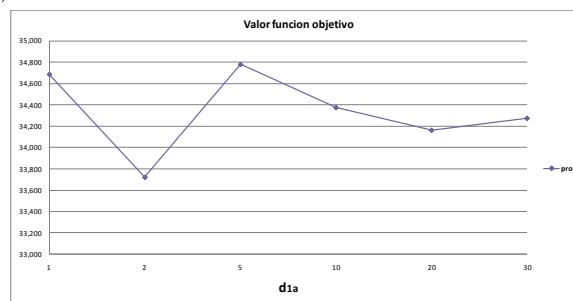
(b)



(c)



(d)



(e)

Figura 30: Tiempo de espera total pasajeros (a); Regulación promedio (b); Tiempo de holding total (c); Tiempo de skipping total (d) y Valor de la función objetivo (e).

Los resultados se presentan en la figura 30. Al observar el tiempo de espera total de pasajeros (figura 30.a), tiende a aumentar levemente con el aumento del tamaño de la ventana. En el caso la regulación promedio (figura 30.b) también se tiende a mantener constante. En relación a las acciones de control, la componente tiempo holding total (figura 30.c) tiende a oscilar entre 12,500 y 11,000; y la componente de tiempo skipping total (figura 30.d) también oscila pero entre 450 y 300. Finalmente al observar el valor de la función objetivo (figura 30.e), no se tiene una tendencia clara.

De este análisis no se pueden sacar conclusiones claras, ya que para diferentes valores de d_{1a} se obtienen diferentes comportamientos en los indicadores, por ende, es preferible mantener un valor intermedio para evitar cualquier variación no esperada.

5.5. Aplicación de predictores a red de transporte público

En el capítulo 3 se presentaron dos estrategias de control HPC para un sistema de transporte público donde se analizó su comportamiento para diferentes ajustes. En este capítulo, en las secciones 5.3.1 y 5.3.2, se presentan dos opciones de predictor de demanda los cuales se analizaron para observar su comportamiento. A continuación se presenta la implementación de la estrategia de control con la metodología de solución de selección de paraderos considerando tres opciones de predicción de demanda:

- Sin predictor, es decir, con demanda real.
- Con predictor de demanda heurístico .
- Con predictor basado en NN.

Como se ha indicado previamente, el objetivo de disponer de un predictor más preciso es mejorar el rendimiento de las estrategias de control aplicadas a una red de transporte público. A continuación se simula la red de transporte (30 réplicas) aplicándole la estrategia de control HPC para observar si se tiene un impacto real en disponer de predictores más precisos.

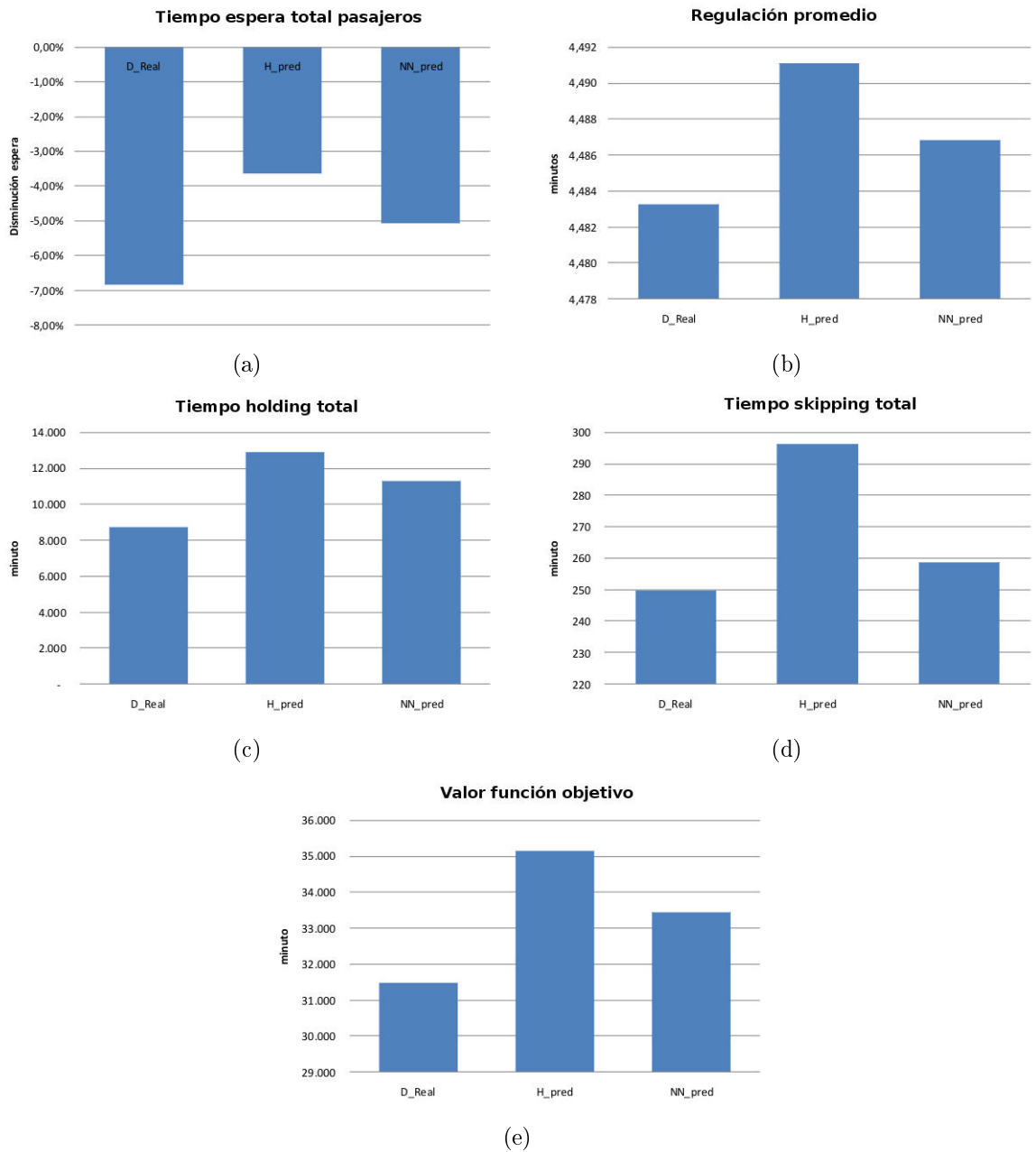


Figura 31: Variación tiempo de espera pasajeros (a); Regulación promedio (b); Tiempo de holding total (c); Tiempo de skipping total (d) y Valor de la función objetivo (e).

En la Figura 31, los gráficos muestran una clara mejoría de los resultados obtenidos al implementar el predictor basado en NN sobre el predictor heurístico; se obtienen mejores tiempos de espera total de pasajeros, un menor nivel de regulación promedio, una menor cantidad de tiempo de holding y skipping total y una función objetivo de menor valor. Esto permite concluir que destinar recursos a mejorar el predictor sí tiene consecuencias positivas en los indicadores y se podría seguir mejorando el predictor buscando llegar al resultado

óptimo que se obtiene considerando la demanda real.

5.6. Discusión

En este capítulo se han presentado una serie de elementos asociados a la predicción de demanda para ser utilizados en estrategia de control HPC para sistemas de transporte público. Se ha buscado proyectar las investigaciones realizadas en predicción de tráfico hacia control de sistemas de transporte público, sobre todo el uso de la herramientas de redes neuronales y generar procesos paralelos para no diseñar un gran predictor que abarque todas las situaciones y paraderos de la red, sino que se enfoque a un desarrollo local asociado a paraderos.

Es importante recalcar en una primera instancia el cambio de enfoque en que se enfrenta la predicción, trabajando directamente con la cantidad de pasajeros que llegan en un intervalo de tiempo y no con instantes de tiempo de llegada de pasajeros. Esto permite evitar tener predictores de alto costo computacional y enfocarse a tener una mayor precisión en la variable que utiliza directamente el controlador.

En relación al punto anterior, se propone utilizar mediciones de errores de predicción directo y no evaluar el desempeño del predictor a través de los indicadores de desempeño del controlador, lo cual facilita el análisis y permite observar directamente el impacto del predictor sin mezclar su rendimiento con el del controlador.

Además se realiza un ajuste de los parámetros del algoritmo de predicción original (heurístico) para observar si la variación de sus parámetros permite obtener mejores resultados de predicción.

Finalmente se concluye que el disponer de un predictor basado en redes neuronales genera mejores resultados, tanto en la medición de su error directo como en los indicadores de desempeño al utilizarlo en conjunto con el controlador. Se logra una adaptación más directa para predecir una serie de datos que son no homogéneos y que dependen fuertemente de los instantes de tiempo.

6. Conclusiones y trabajos futuros

En este trabajo de tesis se ha analizado la aplicación de algoritmos de control predictivo híbrido para sistemas de transporte público.

Se presenta una nueva metodología para utilizar los datos recolectados de un corredor real, facilitados por Transantiago S.A., y poder construir una demanda de pasajeros en paraderos confiable que represente en forma más precisa la realidad, y no construir la demanda a través de modelos de Poisson.

El trabajar con corredores de mayores dimensiones implica proponer nuevas metodologías de resolución para la estrategia de control HPC para que su aplicación en tiempo real sea factible con un tiempo computacional manejable. Esta nueva metodología se basa en la selección de paraderos, que permite trabajar con redes de transporte de gran tamaño. La selección de paraderos permite aplicar las decisiones de control en forma local y no considerando la totalidad del sistema, lo que en redes de transporte de gran tamaño y distancias amplias, no genera mayores distorsiones.

Este nuevo enfoque permite aumentar los pasos hacia futuro del controlador manteniendo el valor de la función objetivo, pero disminuyendo el tiempo de espera y el valor de la regulación. En sistemas de transporte, la demanda no tiene un comportamiento homogéneo entre los días, variando según el día de la semana o en situaciones excepcionales, por eso se analiza el comportamiento de la estrategia de control frente a diferentes tipos de demanda. Este estudio recalca la necesidad de que la selección de la estrategia de control y su configuración son primordiales para obtener resultados satisfactorios. El método de solución del controlador predictivo híbrido basado en estrategia de selección de paraderos se acomoda al funcionamiento de diferentes escenarios. Para algunos escenarios funciona mejor, esto sucede cuando el headway de diseño es pequeño, los buses no llegan a su capacidad máxima o el mayor volumen de demanda se encuentra al final del corredor, en estos casos los resultados son sobresalientes.

Para obtener mejoras con la aplicación de estrategias de control HPC en sistemas de transporte, se requiere disponer de un algoritmo de predicción preciso y que tenga la

capacidad de trabajar con la demanda propuesta real de Transantiago S.A. En trabajos anteriores, solamente se dispone de un predictor ajustado a demandas generadas con el modelo de Poisson. En el capítulo 5 se presenta esta nueva propuesta de predicción, que consiste en un cambio de enfoque en la metodología, ya que no se busca predecir instantes de llegada de pasajeros, sino que enfocarse en cantidades de pasajeros que arriban en intervalos de tiempo. Este cambio de enfoque se basa en que el controlador trabaja con cantidades de pasajeros en intervalos de tiempo y no instantes de llegada de pasajeros, por lo que el algoritmo predictor se puede ajustar adecuadamente a los requerimientos de la estrategia de control.

Para poder comparar directamente los predictores, se propone utilizar la medición de error de predicción directo y no evaluar el desempeño del predictor a través de los indicadores de desempeño del controlador. De esta forma se disocia la medición del rendimiento de estos dos algoritmos permitiendo optimizarlos independientemente y buscar la mejora de los resultados en forma más precisa.

El disponer de un predictor basado en redes neuronales, permite mejores resultados, tanto en la medición de su error directo como en los indicadores de desempeño al utilizarlo en conjunto con el controlador.

A continuación se indican posibles trabajos futuros que se derivan de esta tesis, que son relevantes en la aplicación de estrategias control HPC en sistemas de transporte público.

En relación a la construcción de demandas de pasajeros más representativas de la realidad, se podría mejorar lo que es la determinación de los destinos de pasajeros utilizando la información disponible de Transantiago S.A.

Otra área de mejoras está asociada a implementar nuevas representaciones de sistemas de transporte público no diseñando redes circulares, sino que lineales con dos paraderos terminales y donde los paraderos en un sentido coinciden con paraderos en el sentido contrario. Esto sería de mucha utilidad en la comprensión de sistemas de transporte enmallados, donde por su tamaño, la aplicación de recorridos circulares no es lo más óptimo. Asociado a este punto se debe considerar la aplicación de estrategias de control donde el sistema de transporte

tiene paraderos de transbordo que intercomunica dos recorridos.

Con respecto a la estrategia de control HPC, las mejoras van en la línea de analizar otras acciones de control, como qué tiempos de *holding* puedan ser continuos y no discretos, o la inyección de buses donde los costos operacionales extras se hacen relevantes en relación al costo extra de la flota y la operación. En este caso, la función objetivo debería incluir otros términos asociados a los costos operacionales.

Finalmente el bloque predictor necesario para la aplicación del control HPC, se puede mejorar buscando algoritmos de predicción que consideren el historial no solamente del paradero consultado, sino que también utilicen la información de sus vecinos.

Anexos

Predicción del instante de llegada de usuarios al paradero

Esta metodología de predicción se basa en Sáez (2007), la predicción se alimenta de dos componentes:

- La información proporcionada por los datos en tiempo real disponibles a medida que el tiempo transcurre.
- La información histórica de días anteriores.

De lo anterior se obtiene la predicción del próximo instante de llegada de un usuario a un paradero específico. Para ello, específicamente se calcula el incremento temporal en el cual llegará otro usuario desde el último usuario arribado.

Se denota el incremento en tiempo real como $\overline{\Delta}_{on-line}$ y el incremento *off-line* como $\overline{\overline{\Delta}}_{data-hist}$. Siendo y_m el instante de llegada del usuario m al paradero que se analiza, se describe la hora de llegada predicha del nuevo usuario \hat{y}_{m+1} al paradero como

$$\hat{y}_{m+1} = y_m + \alpha \cdot (\overline{\Delta}_{on-line}) + (1 - \alpha) \cdot (\overline{\overline{\Delta}}_{data-hist}) \quad (18)$$

$\alpha \in [0, 1]$ es un ponderador que otorga pesos distintos a los datos en tiempo real y los datos históricos. Este valor de α dependerá de la confiabilidad de los datos en tiempo real.

Si no se dispone de datos en tiempo real suficientes o representativos del sistema, se estima la hora de llegada con $\alpha = 0$.

Cálculo de incrementos con datos en tiempo real, $\overline{\Delta}_{on-line}$

Como notación se tiene y_m que corresponde al último instante de llegada registrada de un pasajero que es previo al instante actual del sistema $y(t)$.

Además dada una ventana de tiempo δ_{1a} se determina un conjunto de datos que serán los considerados como datos en tiempo real. Este conjunto, corresponde a las horas de llegadas de la parada analizada que se encuentran dentro de la ventana de tiempo de largo δ_{1a} medido

desde $y(t)$ hacia atrás. Estos son todos los datos registrados entre $y(t)$ e $y(t) - \delta_{1a}$. Las anteriores variables se aprecian en la tabla 22

Hora de llegada	Destino
y_1	d_1
y_2	d_2
\vdots	\vdots
$y_{m-\gamma-1}$	$d_{m-\gamma-1}$
$y_{m-\gamma}$	$d_{m-\gamma}$
\vdots	\vdots
y_{m-1}	d_{m-1}
y_m	d_m
\hat{y}_{m+1}	\hat{d}_{m+1}

} δ_{1a}

Tabla 22: Formato de los datos online

Para los otros casos se debe cumplir que:

$$y_m \leq y(t) \tag{19}$$

$$y_{m-\gamma-1} < y(t) - \delta_{1a} \leq y_{m-\gamma} \tag{20}$$

γ corresponde al índice para determinar los valores de y y d que quedan dentro de la ventana de tamaño d_{1a} .

Se define el conjunto $Y_{d_{1a}} = (y_{m-\gamma}, y_{m-\gamma+1}, \dots, y_{m-1}, y_m)$ como los datos en la ventana δ_{1a} y sea $r = \text{card}(Y_{\delta_{1a}})$.

Se obtiene la diferencia promedio de llegada de los usuarios al paradero de esta ventana calculando la siguiente expresión:

$$\bar{\Delta}_{on-line} = \frac{y_m - y_{m-\gamma}}{r - 1} \tag{21}$$

Que se obtiene de:

$$\overline{\Delta}_{on-line} = \frac{\sum_{i^*=1}^{r-1} \Delta_1}{r-1} = \frac{\sum_{i=m-\gamma}^{m-1} y_{i+1} - y_i}{r-1} \quad (22)$$

Cálculo de incrementos con registros históricos (datos off-line), $\overline{\Delta}_{data-hist}$

Se define la variable δ_2 como la cantidad de días de los datos históricos que se considera en la estimación. Los datos registrados del día N-1 hacia atrás corresponden a la data off-line, el valor $y_{n_{[\Omega]}}$ es el último dato registrado de instante de llegada del día Ω . Para cada día Ω de data off-line, se tienen $n_{[\Omega]}$ datos registrados de horas de llegada de pasajeros al paradero en cuestión, los días de data histórica no necesariamente tienen igual cantidad de datos.

Se define la variable δ_{1a} que denota el largo de ventana de los datos en tiempo real que comprende los registro entre la hora actual o de consulta $y(t)$ hacia atrás y de manera análoga se define para la data off-line una variable adicional δ_{1b} que denota el largo de ventana hacia adelante. Estas variables se pueden observar en la tabla 23

Día Ω	
Hora de llegada	Destino
$y_{1(\Omega)}$	$d_{1(\Omega)}$
$y_{2(\Omega)}$	$d_{2(\Omega)}$
\vdots	\vdots
$y_{n_{(\Omega)}-\gamma_1-1}$	$d_{n_{(\Omega)}-\gamma_1-1}$
$y_{n_{(\Omega)}-\gamma_1}$	$d_{n_{(\Omega)}-\gamma_1}$
\vdots	\vdots
$y_{m_{(\Omega)}-1}$	$d_{m_{(\Omega)}-1}$
$y_{m_{(\Omega)}}$	$d_{m_{(\Omega)}}$
$y_{m_{(\Omega)}+1}$	$d_{m_{(\Omega)}+1}$
\vdots	\vdots
$y_{n_{(\Omega)}-\gamma_2}$	$d_{n_{(\Omega)}-\gamma_2}$
$y_{n_{(\Omega)}-\gamma_2+1}$	$d_{n_{(\Omega)}-\gamma_2+1}$
\vdots	\vdots
$y_{n_{(\Omega)}-1}$	$d_{n_{(\Omega)}-1}$
$y_{n_{(\Omega)}}$	$d_{n_{(\Omega)}}$

$y(t)$
 \swarrow

Tabla 23: Datos de un día Ω cualquiera

Considerando un día Ω cualquiera de datos se considera la restricción que siempre $y_{n_{(\Omega)}} - y_{1(\Omega)} > \delta_{1a} + \delta_{1b}$ que significa que las ventanas de datos no deben ser más grandes que el tiempo total de datos registrados de horas de llegada. Los días de la data histórica no necesariamente tienen igual cantidad de datos.

Para δ_{1a} los límites se establecen como $y_{m+\gamma_1-1} < y(t) - \delta_{1a} < y_{m+\gamma_1}$ y además $y_m < y(t)$, luego los datos considerados en la ventana vienen dados por todos los valores $y_{i_1} \in (y(t) - \delta_{1a}, y(t)]$ donde $i_1 = m - \gamma_1, m - \gamma_1 + 1, \dots, m$. Con esto se forma el conjunto $Y_{\delta_{1a}} = \{y_{m-\gamma_1}, y_{m-\gamma_1+1}, \dots, y_{m-1}, y_m\}$ y se define $r_1 = \text{card}(Y_{\delta_{1a}})$.

Para δ_{1b} los límites se establecen como $y_{m+\gamma_2} < y(t) + \delta_{1b} < y_{m+\gamma_2+1}$ y además $y_{m+1} < y(t)$, luego los datos considerados en la ventana vienen dados por todos los valores $y_{i_2} \in (y(t), y(t) + \delta_{1b}]$ donde $i_2 = m + 1, m + 2, \dots, m + \gamma_2 + 1, m + \gamma_2$. Con esto se forma el conjunto $Y_{\delta_{1b}} = \{y_{m+1}, y_{m+2}, \dots, y_{m+\gamma_2-1}, y_{m+\gamma_2}\}$ y se define $r_1 = \text{card}(Y_{\delta_{1a}})$.

Se utilizan ventanas δ_{1a} y δ_{1b} razonables, escogidas por el diseñador.

Uniendo estos dos conjuntos se obtiene el conjunto total de datos off-line.

Sea $y_i \in [y(t) - \delta_{1a}, y(t)] \cup (y(t), y(t) + \delta_{1b}] = [y(t) - \delta_{1a}, y(t) + \delta_{1b}]$ donde:

$i = m - \gamma_1, m - \gamma_1 + 1, \dots, m - 1, m, m + 1, m + 2, \dots, m + \gamma_2 - 1, m + \gamma_2$.

Con esto se forma el conjunto

$Y_{\delta_{1a}, \delta_{1b}} = Y_{\delta_{1a}} \cup Y_{\delta_{1b}} = \{y_{m-\gamma_1}, y_{m-\gamma_1+1}, \dots, y_{m-1}, y_m, y_{m+1}, y_{m+2}, \dots, y_{m+\gamma_2-1}, y_{m+\gamma_2}\}$ y se define finalmente $r = \text{card}(Y_{\delta_{1a}, \delta_{1b}})$. Se tiene entonces que $r = r_1 + r_2 = \gamma_1 + \gamma_2 + 1$.

Con esto se obtiene un promedio de las diferencias entre horas de llegada de la ventana calculado como:

$$\bar{\Delta}_{[\Omega]} = \frac{\sum_{i^*=1}^{r-1} \Delta_{i^*}}{r-1} \quad (23)$$

Aplicando que $r = r_1 + r_2 = \gamma_1 + \gamma_2 + 1$ se tiene

$$\bar{\Delta}_{[\Omega]} = \frac{\sum_{i=m-\gamma_1}^{m+\gamma_2-1} y_{i+1} - y_i}{\gamma_2 + \gamma_1} \quad (24)$$

Hasta el momento, el cálculo es para un sólo día Ω de data histórica. Realizando el mismo procedimiento para los δ_2 de data histórica que se consideran, se obtiene una serie de valores $\bar{\Delta}_q$ donde $j = 1, 2, \dots, \delta_2$.

El resultado $\bar{\bar{\Delta}}_{data-hist}$ será un promedio de los valores $\bar{\Delta}_q$. Esto queda representado en la siguiente ecuación.

$$\bar{\bar{\Delta}}_{data-hist} = \frac{\sum_{q=1}^{\delta_2} \bar{\Delta}_q}{\delta_2} \quad q = 1, 2, \dots, \delta_2 \quad (25)$$

Re-escribiendo la ecuación anterior se obtiene finalmente:

$$\bar{\bar{\Delta}}_{data-hist} = \frac{1}{\delta_2} \cdot \sum_{q=1}^{\delta_2} \left(\frac{y_{q, m_q + \gamma_{2q}} - y_{q, m_q - \gamma_{1q}}}{\gamma_{2q} + \gamma_{1q}} \right) \quad q = 1, 2, \dots, \delta_2 \quad (26)$$

Modelo para predicción de destino de pasajero

Los destinos predichos para la cantidad de pasajeros calculada en el paso anterior se obtienen considerando los destinos de los pasajeros tanto en tiempo real como de los días históricos disponibles, a través de histogramas que se utilizan para asignar a los paraderos una probabilidad y se escoge como posible destino el que tenga un mayor valor. Este procedimiento permite predecir un destino probable, por lo que se realiza tantas veces como pasajeros predichos se tengan.

El histograma de datos en tiempo real se formula como un vector de largo p , que tiene para cada paradero la cantidad de pasajeros que han tenido ese destino en el día actual:

$$h_{actual} = [h_1, h_2, \dots, h_p] \quad h_i \in \mathbb{N}$$

Esto se logra definiendo una ventana de largo d desde el tiempo actual de consulta t hacia atrás, donde se contabilizan los destinos de los pasajeros que llegaron en ese intervalo.

Para los histogramas sobre datos históricos $h_{hist} = [h_1, h_2, \dots, h_p] \quad h_i \in \mathbb{N}$ se sigue un procedimiento similar al histograma en tiempo real, pero la ventana definida también considera un largo d hacia adelante desde el tiempo t .

Con los histogramas históricos se genera un vector con el promedio de todos los días Ω , y se le asocia una probabilidad simple a cada paradero $Ph_{prom}(i)$.

Con la finalidad de tomar en cuenta de que la gente que ha llegado durante el día actual provoca que la probabilidad de estos destinos disminuya se define $delta = 1 - \frac{h_{actual}}{h_{hist}}$, este valor mientras mayor es significa que la aparición de ese destino es muy baja comparada a la aparición histórica, además si tiene una mayor cantidad de apariciones en comparación al promedio histórico toma un valor negativo.

Finalmente la probabilidad de los posibles destinos de los pasajeros se calcula ponderando $delta$, $Ph_{prom}(i)$ y una variable aleatoria P_{rand} .

Predicción de variables basado en redes neuronales

La propuesta a presentada en el capítulo 5 considera la utilización de redes neuronales por lo que a continuación se presenta la base teórica de redes neuronales a utilizar.

La predicción de variables con redes neuronales se entiende como una aplicación particular de la identificación de sistemas, que se describe como la tarea de deducir una descripción matemática, es decir, un *modelo* de un sistema dinámico basándose en una serie de mediciones del sistema.

El desarrollar un modelo a través de redes neuronales implica una serie de pasos, que en este caso se han basado en Nørgaard et al. (2000).

Para poder identificar un sistema mediante redes neuronales, y con esto lograr predecir su comportamiento hacia adelante, es recomendable seguir el procedimiento diagramado en la figura (32):

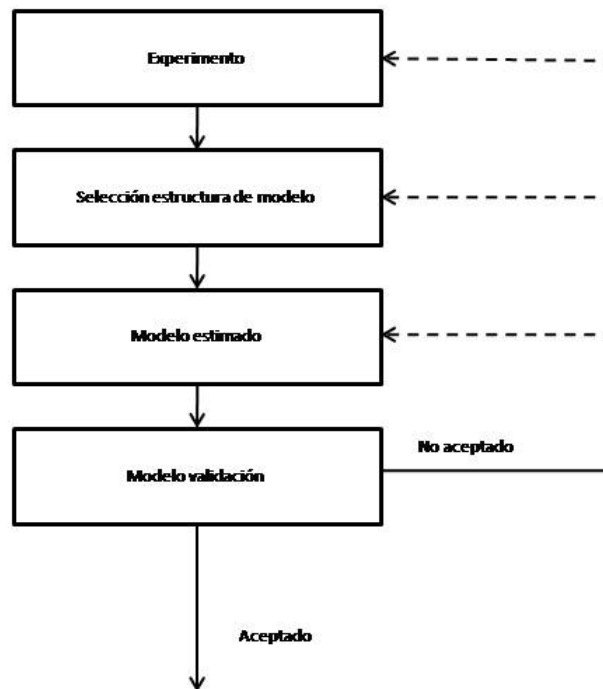


Figura 32: Procedimiento básico de identificación de estructura

Este procedimiento consta de una serie de pasos que se explican a continuación:

1. **Recolección de datos:** El propósito es generar un set de datos que describan cómo el sistema se comporta para el rango completo de operación. La idea es variar la entrada y observar el impacto en la salida como se aprecia en la figura 33.

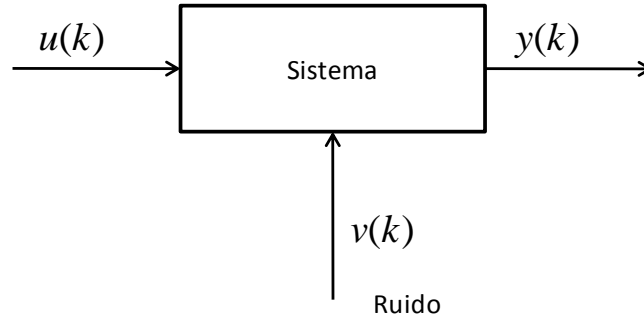


Figura 33: Una entrada es aplicada a el sistema y se observa cómo la salida es afectada

De aquí se obtienen los pares

$$Z^N = \{[u(k), y(k)] | k = 1, \dots, N\} \quad (27)$$

que son utilizados para inferir un modelo para el sistema

2. **Selección de estructura:** Se refiere a seleccionar un conjunto de modelos candidatos donde se debería buscar el modelo definitivo. La selección se separa en dos componentes:
 - a) Seleccionar una “familia” de estructuras modelos consideradas apropiadas para describir el sistema, ejemplos, *estructuras de modelo lineal, red de percepción multicapa (MLP), red de función radial base, etc.*
 - b) Seleccionar un subgrupo de la familia escogida, por ejemplo, en la familia de estructuras lineales, se puede tener una estructura de modelo ARX.
3. **Estimación de parámetros:** después de seleccionar el modelo más adecuado a las características del sistema y los datos disponibles, se realiza una estimación de los parámetros. Típicamente se escoge el modelo que se comporte de mejor forma de acuerdo

a algún criterio. Esta regla debería estar relacionada con el uso pensado del modelo. Generalmente se utiliza como criterio quien tenga el menor error medio cuadrático (RMS) entre las observaciones y lo predicho. Esta fase se conoce generalmente como entrenamiento de la red en redes neuronales.

4. **Validación del modelo:** Cuando un modelo ha sido entrenado, se debe evaluar si se satisfacen los requerimientos necesarios. La validación es cercana al uso pensado del modelo.

En la ejecución de este procedimiento, es muy difícil llegar al modelo idóneo realizando estos pasos una sola vez, por ende se propone ejecutar el procedimiento en forma iterativa, ya que puede ser necesario volver para determinar un número diferente de modelos (paso 2), poner a prueba estructuras de modelos (paso 3), y en el peor caso, inclusive rehacer el experimento (paso 1).

.

Cada uno de los pasos anteriormente expuestos brevemente se explican a continuación con detalle.

Experimento

La realización del experimento es particularmente importante para modelos no lineales. Se debe ser extremadamente cuidadoso en recolectar un set de datos que describan cómo el sistema se comporta en todo el rango de operación.

Un procesamiento inteligente de los datos es generalmente más importante que intentar una gran cantidad de diferentes estructuras de modelo y procesos de entrenamiento. Se busca extraer la información más valorable desde los datos medidos. Para lo cual se pueden realizar una serie de tratamientos:

- **Filtrar:** Se utiliza ampliamente para remover ruido de la medición, distorsiones periódicas, off-set y efectos no interesantes de la dinámica.
- **Eliminar redundancia y outliers:** Algunas veces una gran cantidad de pares entrada-

salida desde un pequeño régimen del rango completo de operación domina el set de datos. Cuando se entrena con este set, se obtiene una alta precisión en el régimen sobrerrepresentado y un bajo desempeño en los otros regímenes. Por lo anterior se deben eliminar los datos redundantes para obtener una incidencia equitativa de la información. Se recomienda además eliminar los outliers ya que generalmente tienen un impacto fatal en los modelos de entrenamiento.

- **Escalar:** Se recomienda eliminar la media y escalar toda la señal a la misma varianza, ya que las señales de mayor magnitud serán dominantes. De esta forma se logra un algoritmo más robusto.

Proceso identificación del sistema

Un sistema dinámico se puede entender como una función S con ciertas entradas y con un vector de salida, tal como se aprecia en la figura 34

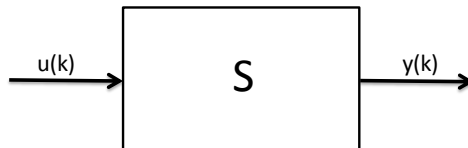


Figura 34: Modelo sistema dinámico

La ecuación que representa el sistema real es:

$$y(k) = S [y(k-1), \dots, y(k-n), u(k-1), \dots, u(k-m)]$$

Donde $y(k)$ es la salida del sistema en el instante actual; $y(k-1), \dots, y(k-n)$ son la salida del sistema para n instantes anteriores y $u(k-1), \dots, u(k-m)$ son la entrada al sistema para m instantes anteriores.

Este sistema se puede modelar mediante estructuras no lineales y en Haykin (1998) se indica que una red de perceptrón multicapa es buena opción para aprender relaciones no lineales desde un set de datos. La estructura del modelo se obtiene de :

$$y(k) = g[\varphi(k), \theta] + e(k)$$

o del predictor de la forma

$$\hat{y}(k|\theta) = g[\varphi(k), \theta] \quad (28)$$

θ es el vector que contiene los parámetros ajustables en la red neuronal conocidos como *pesos* y $\varphi(k)$ es el vector regresor compuesto por:

La estructura del predictor en la ecuación 28 varía según cómo se defina el vector $\varphi(k)$; si se escoje un modelo autoregresivo de entrada externa (ARX) la estructura del modelo es llamado NNARX.

En un modelo NNARX se trabajar con un vector autoregresor compuesto por las siguientes variables:

$$\varphi(k) = [y(k-1), \dots, y(k-n), u(k-d), \dots, u(k-d-m)]^T$$

Es decir, considera un entrada externa u entre los instantes d y m , además de n instantes hacia atrás del vector real de salida $y(k)$.

El modelo general de dos capas utilizado para estimar al sistema real dinámico mediante redes neuronales se puede representar como:

$$\begin{aligned} \hat{y}(k|\theta) &= g[\varphi(k), \theta] \\ &= F_i \left(\sum_{j=1}^{nh} W_{ij} h_{ij}(\theta) + W_{i0} \right) \\ &= F_i \left[\sum_{j=1}^{nh} W_{ij} f_j \left(\sum_{l=1}^{n\varphi} w_{jl} \varphi_l(k) + w_{j,0} \right) + W_{i0} \right] \end{aligned} \quad (29)$$

En la ecuación θ está compuesto por W_{ij} y w_{ij} .

La visión gráfica de la ecuación (29) es la figura 35

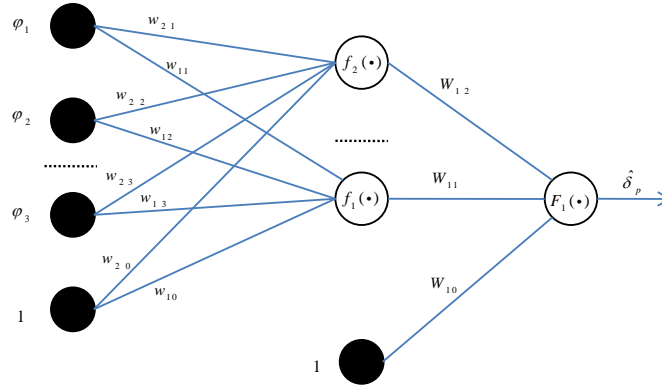


Figura 35: Red Neuronal

Para estimar el vector θ se utiliza el set de entrenamiento definido en la ecuación 27, es decir, el objetivo del set de entrenamiento es ajustar los pesos aproximados de la red, $Z^N \rightarrow \hat{\theta}$. El entrenamiento se detalla en el capítulo 6.

Selección de estructura Además de definir la estructura modelo, es necesario determinar con cuántos datos históricos se desea trabajar, estos no pueden ser ni muy pocos (ya que no se logra modelar la esencia dinámica) ni muchos (se manifiestan como factores comunes en la función de transferencia, pero eso no implica que se genere un problema). Por lo que se trabaja con la metodología de coeficientes de Lipschitz, la cual busca determinar cuánta información pasada es suficiente para generar un modelo de predicción (en 6).

Determinación de Pesos

Como se indicó al inicio de este capítulo, para determinar el valor de los pesos dado un set de datos $Z^N = \{[u(k), y(k)], k = 1, \dots, N\}$ y un modelo candidato $y(k) = \hat{y}(k|\theta) + e(k) = g[k, \theta] + e(k)$, se busca descubrir $Z^N \rightarrow \hat{\theta}$.

Para ver cómo se comporta el modelo con respecto a la salida del sistema real, se estudia el error aproximado de predicción basándose en la introducción de una medida de proximidad en términos del criterio error medio cuadrático (RME).

$$V_N(\theta, Z^N) = \frac{1}{2N} \sum_{k=1}^N [y(k) - \hat{y}(k|\theta)]^2 = \frac{1}{2N} \sum_{k=1}^N \epsilon^2(k, \theta) \quad (30)$$

RME es la mejor medición de error si la distribución del ruido es gaussiano.

Este error se utiliza para determinar los pesos aproximados $\hat{\theta}$ mediante la resolución de:

$$\hat{\theta} = \underset{\theta}{\operatorname{arg\,min}} V_N(\theta, Z^N) \quad (31)$$

Lo que se busca es obtener la ecuación (31) a través de la linealización de $V_N(\theta, Z^N)$:

$$V_N(\theta, Z^N) = V_N(\theta^*, Z^N) + (\theta - \theta^*)^T V'_N(\theta^*, Z^N) + \frac{1}{2} (\theta - \theta^*)^T V''_N(\theta^*, Z^N)$$

Se define gradiente como:

$$G(\theta^*) = V'_N(\theta^*, Z^N) = \left. \frac{\delta V_N(\theta, Z^N)}{\delta \theta} \right|_{\theta=\theta^*}$$

y el Hessiano:

$$H(\theta^*) = V''_N(\theta^*, Z^N) = \left. \frac{\delta^2 V_N(\theta, Z^N)}{\delta \theta^2} \right|_{\theta=\theta^*}$$

Entonces por optimización se sabe que se logra el mínimo $\theta = \theta^*$ cuando $G(\theta^*) = 0$ y $V^T H(\theta^*) V > 0$.

Una forma de enfrentarlo es a través de esquemas iterativos, con lo cual la estimación se realiza como:

$$\theta^{(i+1)} = \theta^{(i)} + \mu^{(i)} f^{(i)} \quad (32)$$

Donde $\theta^{(i)}$ especifica la iteración actual, $f^{(i)}$ la dirección de evolución y $\mu^{(i)}$ representa el tamaño del paso. Las iteraciones se desarrollan hasta que $\theta^{(i)}$ esté lo suficientemente cerca del mínimo $\hat{\theta}$.

Existen diferentes tipos de entrenamiento para trabajar esquemas iterativos. Comúnmente se utiliza el algoritmo back-propagation por su simplicidad, pero no posee buenas características

de convergencia Nørgaard et al. (2000) . Otro método de entrenamiento es el Levenberg-Marquardt. Este método es el estándar para trabajar con minimización del criterio de error media cuadrática, porque posee una rápida convergencia y robustez (Nørgaard et al., 2000). Finalmente se presenta el método Recursivo Gauss-Newton que se utiliza cuando los datos antiguos se van haciendo obsoletos si el sistema es variable en el tiempo.

Sobre estos métodos se profundiza a continuación.

Método Levenbert-Marquardt La dirección de búsqueda $f^{(i)}$ de algoritmos iterativos no tienen que ser necesariamente óptima. Para este método, la dirección se determina vía una aproximación del criterio $L^{(i)}(\theta)$.

Este criterio se basa en el método de Gauss-Newton, que toma ventaja de una aproximación lineal del error,

$$\begin{aligned}\tilde{\epsilon}(k, \theta) &= \epsilon(k, \theta^{(i)}) + [\epsilon'(k, \theta^{(i)})]^T (\theta - \theta^{(i)}) \\ &\quad \epsilon(k, \theta^{(i)}) + [\psi'(k, \theta^{(i)})]^T (\theta - \theta^{(i)})\end{aligned}$$

con lo cual se modifica el criterio de error,

$$V_N(\theta, Z^N) \simeq L^{(i)}(\theta) = \frac{1}{2N} \sum_{k=1}^N \tilde{\epsilon}^2(k, \theta)$$

y de esa forma el gradiente es el mismo que en el método de Newton

$$G(\theta^{(i)}) = \left. \frac{dL^{(i)}(\theta)}{d\theta} \right|_{\theta=\theta^{(i)}} = \frac{1}{N} \sum_{k=1}^N \psi(k, \theta^{(i)}) [y(k) - \hat{y}(k|\theta^{(i)})]$$

pero el Hessiano es distinto

$$R(\theta^{(i)}) = \left. \frac{d^2 L^{(i)}(\theta)}{d\theta^2} \right|_{\theta=\theta^{(i)}} = \frac{1}{N} \sum_{k=1}^N \psi(k, \theta^{(i)}) \psi^T(k, \theta^{(i)})$$

y se denomina Hessiano Gauss Newton.

Considerando lo anterior, el método de Levenberg-Marquardt determina la dirección de búsqueda por medio de una aproximación de $L^{(i)}(\theta)$ que se espera que sea válido solamente en una vecindad alrededor de la iteración actual. si el mínimo de $L^{(i)}(\theta)$ está lejos de la actual

iteración $\theta^{(i)}$, se obtiene una pobre búsqueda, por eso mejor buscar el mínimo de $L^{(i)}(\theta)$ en una vecindad $\delta^{(i)}$.

$$\theta^{(i+1)} = \arg \min L^{(i)}(\theta) \text{ sujeto a } |\theta - \theta^{(i)}| \leq \delta^{(i)}$$

$$\theta^{(i+1)} = \theta^{(i)} + f^{(i)}$$

$$[R(\theta^{(i)}) + \lambda^{(i)}I] f^{(i)} = -G(\theta^{(i)})$$

El algoritmo anterior está diseñado específicamente para mínimos cuadrados no lineales, para otros, $\delta^{(i)}$ es distinto. Existe una relación monótonica entre $\lambda^{(i)}$ y $\delta^{(i)}$ pero generalmente no es fácil de encontrar.

La vecindad del radio $\delta^{(i)}$ se interpreta como la región alrededor de $\theta^{(i)}$ dentro de la cual la aproximación $L^{(i)}(\theta)$ puede ser creída como una aproximación válida de $V_N(\theta, Z^N)$.

Para estimar $\lambda^{(i)}$

No existe expresión directa para determinar λ dado un radio de la región de confianza.

Una propuesta es de Fletcher (1987) donde la idea es observar la forma en que la reducción en el criterio coincide con la reducción prevista de la aproximación $L^{(i)}(\theta)$ y luego ajustar λ de acuerdo con esto.

Para medir la precisión de aproximación se utiliza

$$r^{(i)} = \frac{V_N(\theta^{(i)}, Z^N) - V_N(\theta^{(i)} + f^{(i)}, Z^N)}{V_N(\theta^{(i)}, Z^N) - L^{(i)}(\theta^{(i)} + f^{(i)})}$$

Si $r^{(i)} \approx 1$ $L^{(i)}(\theta^{(i)} + f)$ es una buena aproximación de V_N y λ debería ser reducido por un factor.

Si $r^{(i)} \ll 1$ o negativo, λ debería crecer.

Algoritmo Recursivo Como se indicó previamente los algoritmos recursivos se utilizan cuando los datos antiguos se van haciendo obsoletos si el sistema es variable en el tiempo. Entonces se requeriría de un entrenamiento en tiempo real, pero como el tiempo de cálculo de

modelos con redes neuronales es alto y se requiere de una gran cantidad de información no se recomienda una aplicación en tiempo real. Eso si para análisis offline si se podría trabajar con algoritmos recursivos. en vez de asumir que el set de datos crece con el tiempo, el algoritmo recursivo se repite sobre los set de datos Z^N de características periódicas. Esto permite tener algunas ventajas:

- La implementación es más simple
- menor consumo de memoria
- Se utiliza la redundancia de los datos en forma más efectiva para tener una rápida convergencia

En particular se presenta el método recursivo de Gauss-Newton.

Método recursivo de Gauss-Newton

Cuando se usa en forma offline, este método asume que el set de datos Z^N un cierto número de veces $u(k) = u(k + N) = u(k + 2N) = \dots$ y $y(k) = y(k + N) = y(k + 2N) = \dots$

De esta forma el criterio a minimizar cambia a

$$V_k(\theta, Z^k) = \frac{1}{2k} \sum_{j=1}^k \epsilon^2(j, \theta)$$

por lo que la actualización de Gauss-Newton está dada por

$$\theta(k) = \theta(k - 1) - R^{-1}(k) V'_k(\theta(k - 1), Z^k)$$

Y el gradiente toma la forma de

$$V'_k(\theta, Z^k) = -\frac{1}{k} \sum_{j=1}^k \psi(j, \theta) \epsilon(j, \theta) = \frac{k-1}{k} V'_{k-1}(\theta, Z^{k-1}) - \frac{1}{k} \psi(k, \theta) \epsilon(k, \theta)$$

asumiendo que $\theta(k - 1)$ minimiza el criterio en el evento $k - 1$ se tiene que

$$V'_{k-1}(\theta, Z^k) = -\frac{1}{k} \psi(k, \theta(k - 1)) \epsilon(k, \theta(k - 1))$$

El Hessiano Gauss-Newton se reescribe como

$$R(k) = \frac{1}{k} \sum_{j=1}^k \psi(j, \theta) \psi^T(j, \theta) \\ R(k-1) + \frac{1}{k} [\psi(k, \theta) \psi^T(k, \theta) - R(k-1)]$$

Definiendo $\psi(k) \equiv \psi(k, \theta(k-1))$ se puede trabajar con una notación simplificada

$$\theta(k) = \theta(k-1) + \frac{1}{k} R^{-1} \psi(k) [y(k) - \hat{y}(k|(k-1))]$$

$$R(k) = R(k-1) + \frac{1}{k} [\psi(k) \psi^T(k) - R(k-1)]$$

Algoritmo de Lipschitz

1. Para un largo determinado de delay, calcular el cociente de Lipschitz

$$q_{ij} = \left| \frac{y(t_i) - y(t_j)}{\varphi(t_i) - \varphi(t_j)} \right| \quad i \neq j$$

Donde $\|\cdot\|$ representa la norma Euclidiana, $y(t)$ salida del sistema y $\varphi(t)$ entrada del sistema.

2. Seleccionar los p cocientes más grandes $p=0,01N \sim 0,02N$
3. Calcular el criterio $\bar{q}(n) = \left(\prod \sqrt[n]{q^{(n)}(k)} \right)^{\frac{1}{p}}$
4. Repetir cálculo anterior para diferentes largos de delay.
5. Graficar el criterio en función de la ventana de delay y escoger el óptimo número de regresores como el punto de inflexión.

Referencias

- Adamski, A. (1996). *Flexible dispatching control tools in public transport*. Heidelberg, Germany: Springer.
- Adamski, A. & A. Turnau (1998). Simulation support tool for real-time dispatching control in public transport. *Transportation Research -Part A* 32(2), 73–87.
- Barnett, A. (1974). On controlling randomness in transit operations. *Transportation Science* (8), 102–116.
- Bell, M. G. H. (1993). The estimation of an origin-destination matrix from traffic counts. *Transportation Science* 17(2), 198–217.
- Bemporad, A. & M. Morari (1999). Control of systems integrating logic, dynamics and constraints. *Automatica* 35, 407–427.
- Casas, J. (2004). *Estimation of the Transport Demand for Real-Time Applications*. Ph. D. thesis, Universitat Politècnica de Catalunya.
- Cascetta, E. (1984). Estimation of trip matrices from traffic counts and survey data: a generalised least square estimator. *Transportation Research* 18B, 289–299.
- Eberlein, X. (1995). *Real-Time Control Strategies in Transit Operations: Models and Analysis*. PhD thesis. Ph. D. thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Eberlein, X., N. Wilson, & D. Bernstein (1999). Modeling real-time control strategies in public transit operations. *Lecture Notes in Economics and Mathematical Systems* 471, 325–346.
- Eberlein, X., N. Wilson, & D. Bernstein (2001). The holding problem with real-time information available. *Transportation Science* 35(1), 1–18.

- Fisk, C. S. (1989). Trip matrix estimation from link traffic counts: The congested network case. *Transportation Research 23B*, 331–336.
- Fletcher, R. (1987). *Practical Methods of Optimization* (2nd edition ed.). Wiley & Sons.
- Florian, M. & Y. Chen (1993). A coordinate descent method for the bilevel o-d matrix adjustment problem. *Publication, Centre de Recherche sur les Transports, Université de Montréal*.
- Fu, L. & Q. Liu (2003). A real-time optimization model for dynamic scheduling of transit operations. *scheduling of transit operations. 82nd Annual Meeting of the Transportation Research Board, Washington, DC*.
- Furth, P. (1986). Zonal route design for transit corridors. *Transportation Science 20*(1), 1–12.
- Haykin, S. (1998). *Neural Networks: A comprehensive Foundation* (2nd edition ed.). Prentice Hall.
- Hegyi, A. (2004). *Model Predictive Control for Integrating Traffic Control Measures. PhD Thesis*. Ph. D. thesis, Delft University of Technology, Delft Center for Systems and Control.
- Hegyi, A., B. De Schutter, & H. Hellendoorn (2005). Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C 13*, 185–209.
- Hickman, M. (2001). An analytic stochastic model for the transit vehicle holding problem. *Transportation Science 35*(3), 215–237.
- Jordan, W. & M. Turnquist (1979). Zone scheduling of bus routes to improve service reliability. *Transportation Science 13*(1).
- Karar, G., G. Mušič, I. Škrjanc, & B. Zupančič (2007a). Hybrid fuzzy modeling for model predictive control. *Journal of Intelligent and Robotic Systems 50*, 297–319.

- Karer, G., I. Škrjanc, B. Zupančič, & G. Mušič (2007b). Hybrid fuzzy model-based predictive control of temperature in a batch reactor. *Computers and Chemical Engineering* 31, 1552–1564.
- Lam, W., J. Zhou, & Z. Sheng (2002). A capacity restraint transit assignment with elastic line frequency. *Transportation Research* 36B, 919–938.
- Leblanc, L. & K. Farhangian (1982). Selection of a trip table which reproduces observed link flows. *Transportation Research* 16B, 83–88.
- Lin, G., P. Liang, P. Schonfeld, & R. Larson (1995). Adaptive control of transit operations, final report, md-26-7002. *US Department of Transportation*.
- Maher, M. J. (1985). The análisis of partial registration-plate data. *Traffic Engineering and Control* 26, 495–497.
- Mahmassani, H. S. & X. Zhou (2006). Dynamic od demand estimation using automatic vehicle identification data. *IEEE Transactions on Intelligent Transportation Systems* 7(1), 105–114.
- Nguyen, S. (1977). Estimating an od matriz from network data: A network equilibrium approach. *Centre de Recherche sur les Transports, Universite de Montreal, Montreal, Québec* (87).
- Nørgaard, M., O. Ravn, N. Poulsen, & L. Hansen (2000). *Neural Networks for Modelling and Control of Dynamic Systems*. London: Springer-Verlag.
- Ortúzar, J. (1983). Nested logit models for mixed-mode travel in urban corridors. *Transportation Research* 17A(4), 283–299.
- Osuna, E. & G. Newell (1972). Control strategies for an idealized public transportation system. *Transportation Science* 6, 52–72.
- Rajbhandari, R., S. Chien, & J. Daniel (2003). Estimation of bus dwell times with automatic passenger counter information. *Transportation research record* (1841), 141–148.

- Sáez, D., C. Cortés, F. Milla, A. Nuñez, M. Riquelme, & A. Tirachini (2009). Hybrid predictive control strategy for a public transport system with uncertain demand. en segunda revisión. *Transpormetrica*.
- Sáez, E. (2007). Estrategias de control predictivo híbrido en tiempo real para sistemas de regulación de intervalos en sistemas de transporte público. memoria para optar al título de ingeniero civil electricista. Universidad de Chile.
- Suh, W., K. Chon, S. Rhee, et al. (2002). Effect of skip-stop policy on a korean subway system. *Transportation Research Record 1793*, 33–39.
- Sun, A. & M. Hickman (2004). The holding problem at multiple holding stations [online]. <http://fugazi.engr.arizona.edu/caspt/sun.pdf>.
- Sun, A. & M. Hickman (2005). The real-time stop-skipping problem. *Journal of Intelligent Transportation Systems 9*(2), 91–109.
- Turnquist, M. & S. Blume (1980). Evaluating potential efectiveness of headway control strategies for transit systems. *Transportation Research Record 746*, 25–29.
- Van Zuylen, H. & L. Willumsen (1980). The most likely trip matrix estimated from traffic counts. *Transportation Research 14B*, 281–293.
- Volmuller, I. & R. Hamerslag (Eds.) (1984). *A Systems Dynamics Approach to the Estimation of Entry and Exit O-D Flows*. VNU Science Press.
- Welding, P. (1957). The instability of a close-interval service. *Operational Research Quarterly 8*(3), 133–148.
- Yu, B. & Z. Yang (2007). A dynamic holding strategy in public transit systems with real-time information. *Applied Intelligence 31*(1), 69–80.
- Ziliaskopoulos, A. & S. Peeta (2001). Foundations of dynamic traffic assignment: The past, the present and the future. *Networks and Spatial Economics 1*(3/4), 233–266.

Zolfaghari, S., N. Azizi, & M. Jaber (2004). A model for holding strategy in public transit systems with real-time information. *International Journal of Transport Management* 2, 99–110.