

Raziskava in razvoj empiričnih modelov za napovedovanje odjema toplote v vročevodnem sistemu

Ervin STRMČNIK, Primož POTOČNIK, Edvard GOVEKAR

Izvleček: Napovedovanje rabe energije v sistemu daljinskega ogrevanja je ključnega pomena za premišljeno upravljanje z energetskimi viri in za zagotavljanje trajnostne energetske oskrbe. Raziskali smo, kateri dejavniki najbolj vplivajo na odjem toplote in na kakšne načine lahko izboljšamo točnost napovedovanja. Ugotovili smo, da imajo največji vpliv temperatura ozračja, pretekli odjem toplote in sončno sevanje. Razvili smo devet empiričnih modelov za napovedovanje odjema toplote v vročevodnem sistemu. Kot referenčna modela smo določili model random-walk in model temperaturne korelacije, nato pa smo razvili kompleksnejši regresijski model in več vrst avtoregresijskih modelov. Kakovost empiričnih modelov smo ocenili s pomočjo križnega vrednotenja na podlagi napake napovedi, ki je predstavljala srednjo absolutno normalizirano napako glede na maksimalno prenosno kapaciteto vročevoda. Kvaliteto napovedi odjema toplote smo izboljšali z uvedbo populacijskih značilnik za dneve v tednu in dodatnih spremenljivk za upoštevanje rasti sistema ter sezonskega cikla. Za najkakovostnejši empirični model se je izkazal model iz skupine avtoregresijskih modelov (model SR), ki smo ga razvili s pomočjo metode postopne izbire regresorjev.

Ključne besede: daljinsko ogrevanje, napovedovanje odjema toplote, predikcijski modeli, izpeljava značilnik, izbira regresorjev, regresijski model, avtoregresijski model

■ 1 Uvod

Napovedovanje odjema toplote sodi zaradi prisotnosti kompleksnih procesov med zahtevnejše naloge daljinskega ogrevanja. Kakovostne napovedi odjema toplote v vročevodnem sistemu so zelo pomembne s stališča učinkovite rabe energije, kar v tehničnem smislu zajema učinkovito krmiljenje energetskih sistemov, v ekonomskem smislu pa uskladitev prihodnjih potreb odjemalcev s pogodbeno določenimi odjemi. Aktivnosti, povezane z napovedovanjem odjema toplote, zahtevajo celovito poznavanje problematike daljinskega ogrevanja.

Ervin Strmčnik, mag. inž. str., doc. dr. Primož Potočnik, univ. dipl. inž., prof. dr. Edvard Govekar, univ. dipl. inž., vsi Univerza v Ljubljani, Fakulteta za strojništvo

Napovedovanje rabe energije [1, 2] je predpogoj za uspešno optimizacijo in upravljanje tehničnih procesov, povezanih s proizvodnjo energije in daljinskim ogrevanjem [3]. Daljinsko ogrevanje in premišljeno upravljanje z viri [4, 5] sta ključna elementa učinkovite rabe energije, trajne energetske oskrbe in zmanjšanja emisij toplogrednih plinov. Pri napovedovanju se srečujemo z mnogimi težavami, kot so dnevne konice, letna, tedenska in urna nihanja. Zelo težko je napovedovati hitro spreminjajoče se dogodke. Proučevanja fluktuacij sta se temeljito lotila švedska raziskovalca Gadd in Werner [6], ki poročata o ugotovitvah glede nihanj na letnem, dnevnem in urnem nivoju. Glavni razlog za nihanja na letnem nivoju je temperatura ozračja, medtem ko na odjem na dnevnem in urnem nivoju v veliki meri vpliva obnašanje odjemalcev. Modeli za napovedovanje odjema toplote so zelo podobni modelom

za napovedovanje odjema plina. Njihova formulacija je nazorno opisana na primeru napovedovanja odjema zemeljskega plina [7]. Raziskave in primerjave različnih modelov so bile sistematično opravljene v študiji o napovedovanju izpustov dušikovih oksidov [8]. Ocena kakovosti modelov je možna na osnovi različnih statističnih mer [9, 10], ki kvantitativno ovrednotijo napako napovedi. Hyndman in Koehler [11] pojasnjujeta uporabnost in ustreznost določenih statističnih mer za ocenjevanje napak modelov, med drugim srednjo absolutno napako, normalizirano glede na maksimalno prenosno kapaciteto. Podobna mera je bila uvedena pri vrednotenju sistemov za napovedovanje porabe zemeljskega plina [7]. Normalizacija napak [12] je pomembna za lažjo interpretacijo in primerjavo rezultatov. Na odjem toplote iz distribucijske mreže daljinskega ogrevanja vplivajo številni dejavniki. Pri izbiri vplivnih

dejavnikov je potrebno upoštevati vpliv vremenskih in nevremenskih dejavnikov. Od vremenskih dejavnikov so najpogosteje obravnavani vplivi temperature ozračja T , sončno sevanje S , relativna vlažnost H in hitrosti vetra W . Od vseh zgoraj naštetih vremenskih dejavnikov ima največji vpliv temperatura ozračja [13], medtem ko je veliko težje oceniti vpliv sončnega sevanja, hitrosti vetra [14] in relativne vlažnosti. Poleg vremenskih dejavnikov [15] je v modele za napovedovanje odjema smiselno vključiti tudi obnašanje odjemalcev [13], ki prav tako vpliva na vzorce odjema toplote.

V okviru raziskave smo analizirali, kateri dejavniki vplivajo na odjem toplote v vročevodnem sistemu in na kakšne načine lahko izboljšamo točnost napovedi odjema. Zgradili smo devet empiričnih modelov za napovedovanje odjema toplote v vročevodnem sistemu in jih ocenili na podlagi problematiki primerne statistične mere. Kvaliteto napovedi odjema toplote smo izboljšali z uvedbo populacijskih značilk za dneve v tednu in dodatnih spremenljivk za upoštevanje rasti sistema ter sezonskega cikla. Z metodo postopne izbire regresorjev smo uspeli razviti avtoregresijski model, ki ga odlikujejo velika točnost napovedi, robustnost in nizka kompleksnost glede na število vključenih dejavnikov.

2 Podatki

Pri raziskavi smo uporabili podatke o odjemu toplote na področju

Mestne občine Ljubljana od septembra 2008 do februarja 2013 (slika 1), ki nam jih je dalo podjetje Energetika Ljubljana. Na voljo smo imeli podatke na dnevem nivoju za naslednje fizikalne veličine:

- temperatura ozračja T [°C],
- odjem toplote Q [MW],
- sončno sevanje S [W/m²],
- hitrost vetra W [m/s],
- relativna vlažnost H [%].

Na osnovi analiz in poznavanja problematike smo časovno vrsto podatkov ločili na poletno in zimsko sezono. Pri raziskovanju smo obravnavali zgolj podatke zimske sezone. Obravnavanje poletne sezone je manj aktualno zaradi nizkih vrednosti in manj dinamičnega odjema toplote v poletnih mesecih.

Upoštevanje dodatnih vplivnih dejavnikov, kot sta spreminjanje letnih časov in obnašanje uporabnikov, smo izvedli z uvedbo dodatnih spremenljivk. Linearni čas t_{lin} omogoča upoštevanje rasti distribucijskega sistema. Cikličnost na letni ravni (letni časi) smo vključili v obravnavo z dodatno spremenljivko t_{cos} . Obnašanje uporabnikov in s tem tedensko ponavljajoče se vzorce odjema toplote Q smo zajeli z uvedbo populacijskih značilk za dneve v tednu: d_{WD} , d_{SAT} in d_{SUN} , ki označujejo delovne dni, sobote in nedelje.

3 Pristop k napovedovanju

Empirične modele za napovedovanje odjema toplote smo zgradili na

predpostavki, da je odjem toplote vroče vode enak količini proizvedene toplote vroče vode na izvoru. Privzeli smo, da so pridobljeni podatki zadostovali za izgradnjo predikcijskega modela za napoved odjema toplote v vročevodnem sistemu, s katerim smo želeli napovedati odjem toplote s predikcijskim horizontom enega dne $h = 1$ na nivoju dnevne ločljivosti:

$$Q(t + h), \quad h = 1 \text{ dan} \quad (1)$$

Za ocenjevanje modelov smo vpejali normalizirano napako MARNE (ang. *mean absolute range normalized error*), ki je relativna mera glede na velikost sistema in enostavna za interpretacijo tako v tehničnem kot v ekonomskem smislu. Izračunamo jo kot povprečje absolutnih razlik dejanskih Q_a in napovedanih vrednosti Q_p odjema toplote, ki ga delimo z maksimalno prenosno kapaciteto omrežja Q_{max} :

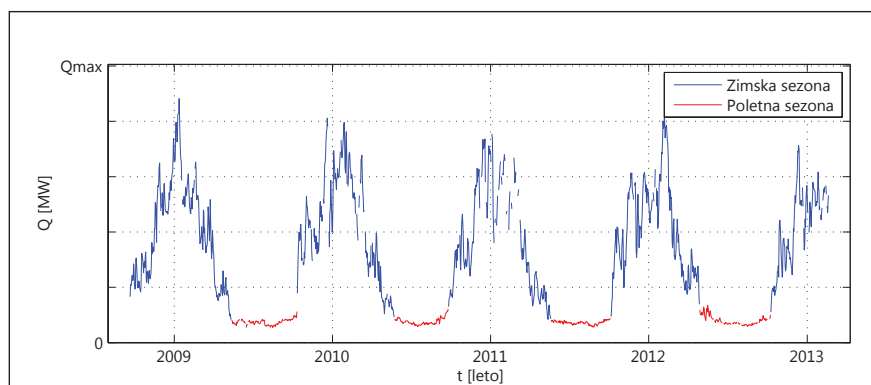
$$\text{MARNE} = 100$$

$$\frac{1}{N_d} \sum_{t=1}^{N_d} |Q_a(t) - Q_p(t)| \frac{100}{Q_{max}}, \quad (2)$$

$$t = 1, 2, \dots, N_d.$$

pri čemer je N_d oznaka za število elementov časovne vrste. MARNE je primerna mera za interpretacijo rezultatov napovedovanja, saj v imenovalcu operira z vrednostjo Q_{max} , ki normira napako napovedi glede na velikost distribucijskega omrežja. Manjša napaka MARNE pomeni boljšo kakovost napovedi modelov za odjem toplote v vročevodnem sistemu.

Pri vrednotenju modelov za napovedovanje odjema toplote smo poleg napake MARNE uporabili tudi Diebold-Marianovo statistiko (DBM) [16], ki med sabo primerja napovedi dveh različnih modelov in ugotavlja, če sta modela statistično različna. Predpogoj za interpretacijo rezultata DBM je izbira intervala zaupanja. Če izberemo stopnjo značilnosti 0,05, to pomeni, da sta modela ob upoštevanju pred-



Slika 1. Odjem toplote Q od septembra 2008 do februarja 2013 na dnevni ravni

postavke o obojestranskem testu statistično značilno različna pri vrednosti $|DBM| > 1,96$.

Posplošitvene sposobnosti modelov za napovedovanje smo ocenili na osnovi križnega vrednotenja modelov, ki zahteva razdelitev razpoložljivih podatkov v dve skupini. Prvo skupino predstavljajo učni podatki, drugo pa testni. Razmerje, ki definira delež učnih in testnih podatkov, je definirano z oznako α . Na osnovi analiz in narave podatkov smo pri razvoju modelov upoštevali vrednost $\alpha = 0,6$, kar pomeni, da je znašal delež učnih podatkov 60 %, medtem ko je bil delež testnih podatkov 40 %. Vse opisane modele smo razvili na osnovi učnih podatkov in jih nato testirali na neodvisnih testnih podatkih. Ključni kriterij za ocenjevanje kakovosti modelov je bila napaka MARNE na testnih podatkih.

■ 4 Opis modelov za napovedovanje

V okviru raziskovalnega dela smo razvili 9 empiričnih modelov za napovedovanje odjema toplote:

- referenčna modela RW (ang. *Random walk*) in TK (model temperaturne korelacije),
- regresijski model REG,
- avtoregresijski modeli ARX1, ARX2, ARX3, ARX4, SR (ang. *Stepwise regression*) in ARIX. Avtoregresijski modeli se med sabo razlikujejo glede izbire vključenih regresorjev. Model SR smo razvili s pomočjo postopne izbire regresorjev.

Nabor vključenih regresorjev za obravnavane modele je prikazan v tabeli 1. Najenostavnejši izmed modelov je bil model RW, ki je predstavljal referenčni model in je služil za primerjavo z drugimi modeli. Model temperaturne korelacije je bil zgrajen na osnovi ugotovljene močne negativne linearne povezanosti med spremenljivkama Q in T . Njun korelacijski koeficient je v zimski sezoni znašal $R = -0,98$. Medtem ko sta enostavna modela RW in TK služila za približno oceno napak napovedanih vrednosti, je bil model REG prvi kompleksnejši model. Vanj je bilo poleg dodatnih značilnik t_{lin} , t_{cos} , d_{WD} , d_{SAT} in d_{SUN} vključenih še pet regresorjev temperature ozračja T (z vložitveno dimenzijo $VD = 5$). Avtoregresijski modeli ARX (ang. *autoregressi-*

ve models with exogenous terms) so poleg avtoregresijskega člena AR vključevali še zunanje vplivne dejavnike X . V našem primeru avtoregresijski del zastopa odjem toplote Q , medtem ko zunanje dejavnike predstavljajo spremenljivke T , S , W in H . Model SR je avtoregresijski model, ki smo ga razvili s pomočjo metode postopne izbire regresorjev, ki nam pomaga določiti statistično vplivne regresorje iz nabora večjega števila potencialno vplivnih dejavnikov. Ideja modela ARIX je bila napovedovanje razlik odjema toplote ΔQ namesto napovedovanja absolutnega odjema toplote Q .

■ 5 Rezultati

Rezultati napovedi po modelih so zbrani v tabeli 2. Tabela prikazuje vrednosti učnih in testnih napak MARNE. S sivim poljem je označena vrednost najmanjše testne napake, ki je končni kriterij ocenjevanja kakovosti modelov. V drugem stolpcu je informativno dodano število v posameznem model vključenih regresorjev.

Rezultati tabele 2 so grafično prikazani tudi na sliki 2. Na abscisni osi

Tabela 1. Pregled uporabljenih regresorjev po modelih

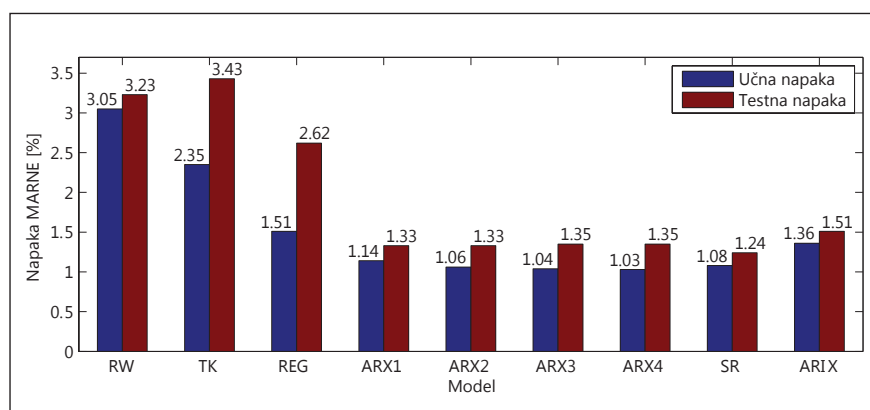
Model	RW	TK	REG	ARX1	ARX2	ARX3	ARX4	SR	ARIX
Napoved	$Q(t+1)$	$Q(t+1)$	$Q(t+1)$	$Q(t+1)$	$Q(t+1)$	$Q(t+1)$	$Q(t+1)$	$Q(t+1)$	$\Delta Q(t+1)$
Regresorji									
Dodatne spremenljivke			t_{lin}	t_{lin}	t_{lin}	t_{lin}	t_{lin}	t_{lin}	t_{lin}
			t_{cos}	t_{cos}	t_{cos}	t_{cos}	t_{cos}	t_{cos}	t_{cos}
			d_{WD}	d_{WD}	d_{WD}	d_{WD}	d_{WD}	d_{WD}	d_{WD}
			d_{SAT}	d_{SAT}	d_{SAT}	d_{SAT}	d_{SAT}	d_{SAT}	d_{SAT}
			d_{SUN}	d_{SUN}	d_{SUN}	d_{SUN}	d_{SUN}	d_{SUN}	d_{SUN}
Temperatura ozračja T		$T(t+1)$	$T(t+1)$	$T(t+1)$	$T(t+1)$	$T(t+1)$	$T(t+1)$	$T(t+1)$	$T(t+1)$
			$T(t)$	$T(t)$	$T(t)$	$T(t)$	$T(t)$	$T(t)$	$T(t)$
			$T(t-1)$	$T(t-1)$	$T(t-1)$	$T(t-1)$	$T(t-1)$		
			$T(t-2)$	$T(t-2)$	$T(t-2)$	$T(t-2)$	$T(t-2)$		
			$T(t-3)$	$T(t-3)$	$T(t-3)$	$T(t-3)$	$T(t-3)$		
Odjem toplote Q	$Q(t)$			$Q(t)$	$Q(t)$	$Q(t)$	$Q(t)$	$Q(t)$	$Q(t)$
				$Q(t-1)$	$Q(t-1)$	$Q(t-1)$	$Q(t-1)$		
				$Q(t-2)$	$Q(t-2)$	$Q(t-2)$	$Q(t-2)$		
				$Q(t-3)$	$Q(t-3)$	$Q(t-3)$	$Q(t-3)$	$Q(t-3)$	$Q(t-3)$
Sončno sevanje S					$S(t+1)$	$S(t+1)$	$S(t+1)$	$S(t+1)$	$S(t+1)$
					$S(t)$	$S(t)$	$S(t)$		
Hitrost vetra W						$W(t+1)$	$W(t+1)$		
Relativna vlažnost H							$H(t+1)$		

Tabela 2. Vrednosti učnih in testnih napak MARNE po modelih

Model	Št. vključenih regresorjev	Učna napaka MARNE [%]	Testna napaka MARNE [%]
RW	1	3,05	3,23
TK	1	2,35	3,43
REG	10	1,51	2,62
ARX1	14	1,14	1,33
ARX2	16	1,06	1,33
ARX3	17	1,04	1,35
ARX4	18	1,03	1,35
SR	10	1,08	1,24
ARIX	10	1,36	1,51

so navedeni vsi obravnavani empirični modeli. Modri stolpci prikazuje vrednost učne napake MARNE, medtem ko rdeči stolpci ponazarjajo testne napake MARNE.

temveč tudi edini statistično različen od vseh preostalih modelov. Za ponazoritev rezultatov sta na *sliki 3* prikazana dejanski in napovedani odjem toplote modela SR za testno

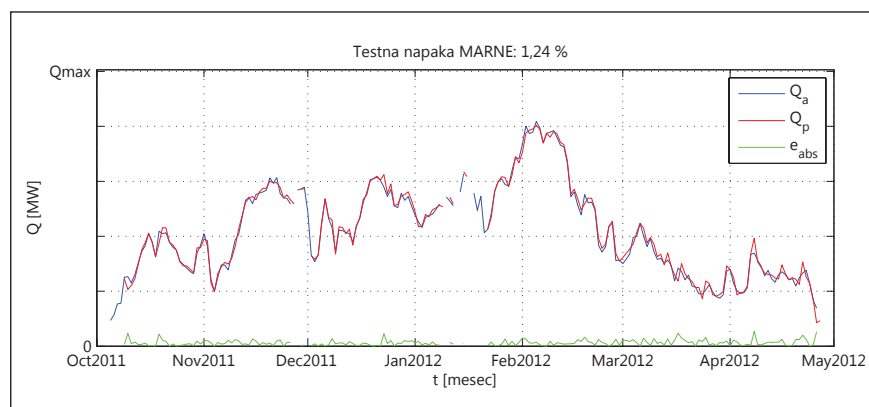
**Slika 2.** Rezultati empiričnih modelov za napovedovanje odjema toplote v vročevodnem sistemu

Iz grafičnega prikaza (*slika 2*) vidimo, da je bila učna napaka referenčnega RW modela 3,05 % in testna 3,23 %. Pri modelu TK na osnovi učne napake in primerjave z modelom RW opazimo, da smo model TK bolje naučili, kajti učna napaka se je zmanjšala za 25 %, vendar je imel model slabše posplošitvene sposobnosti. Testna napaka modela TK se je povečala na vrednost 3,43 %. Najmanjšo učno napako smo dobili z modelom ARX4. Ker smo za osnovno mero kakovosti modela določili testno napako, smo za najkakovostnejši model izbrali model SR, v katerem je vrednost testne napake MARNE znašala 1,24 % in je bila kar 60 % manjša od osnovnega modela RW.

Pri primerjavi modelov s pomočjo Diebold-Marianove statistike se je izkazalo, da model SR ni bil zgolj najkakovostnejši empirični model,

obdobje od oktobra 2011 do maja 2012.

Pri analizi rezultatov časovne vrste najkakovostnejšega modela SR smo opazili veliko boljše prilaganje napovedanih vrednosti Q_p h krivulji dejanskega odjema toplote Q_a kot pri ostalih modelih, kar je vidno tudi na

**Slika 3.** Dejanski in napovedani odjem toplote modela SR

osnovi grafa absolutnih napak, ki so na *sliki 3* prikazane z zeleno barvo.

6 Zaključek

Rezultate raziskave lahko strnemo v sledeče zaključke:

- Med odjemom toplote in temperaturo ozračja obstaja močna linearna povezanost. Najpomembnejši vpliv na kakovost napovedi odjema toplote $Q(t+1)$ predstavlja regresor prihodnjih temperatur ozračja $T(t+1)$.
- Med vremenskimi vplivi je poleg temperature ozračja priporočljivo upoštevati tudi vpliv sončnega sevanja $S(t+1)$, kar je skladno z rezultati na področju raziskav zemeljskega plina [17]. Ugotovili smo, da veter in vlaga nimata pomembnega vpliva na odjem toplote.
- Pri gradnji modelov za napovedovanje odjema toplote je za izboljšanje kakovosti napovedi priporočljivo upoštevanje preteklih odjemov toplote, kar se je izkazalo pri avtoregresijskih modelih.
- Model za napovedovanje smo izboljšali z dodatno uvedeno spremenljivko za upoštevanje rasti distribucijskega sistema t_{lin} , spremenljivko za upoštevanje sezonskega cikla t_{cos} in s populacijskimi značilniki za označevanje delovnih dni, sobot in nedelj d_{WD} , d_{SAT} in d_{SUN} .

Na osnovi testne napake MARNE smo za najkakovostnejši empirični model za napovedovanje odjema toplote izbrali model SR, ki predstavlja največji prispevek opravljene raziskave. V model so bili vključeni sledeči regresorji: t_{lin} , t_{cos} , d_{WD} , d_{SAT}

d_{SUN} , $T(t+1)$, $T(t)$, $Q(t)$, $Q(t-3)$ in $S(t+1)$, ki so se med raziskavami izkazali za statistično vplivne dejavnike na odjem toplote iz distribucijskega omrežja. Z Diebold-Marianovo statistiko smo ugotovili, da je bil model SR tudi statistično različen od vseh preostalih modelov. Zaradi manjšega števila vključenih regresorjev glede na ostale modele iz družine modelov ARX, ki so se na podlagi napake MARNE izkazali za kakovostne, je model SR izkazal veliko mero robustnosti in točnosti napovedovanja in je dovolj enostaven, da bi bil primeren za uporabo na realnih industrijskih aplikacijah. Zaradi visoke linearne povezanosti med odjemom toplote Q in temperaturo ozračja T ne pričakujemo bistvenih izboljšav z uporabo nelinearnih modelov, kar je skladno z rezultati na področju napovedovanja odjema plina [7, 17].

Literatura

- [1] Yeo, I. A., Yoon, S. H., Yee, J. J.: Development of an urban demand forecasting system to support environmentally friendly urban planning, *Applied Energy* 110, 304–317 (2013).
- [2] Thaler, M., Poredoš, A., Grabec, I.: Napoved rabe energije z uporabo neparametrične regresije in genetskega algoritma, diplomsko delo; Univerza v Ljubljani, Fakulteta za strojništvo, 2004.
- [3] Thaler, M.: Analitično-empirični model odjema toplote iz daljinskega ogrevalnega telesa, doktorsko delo; Univerza v Ljubljani, Fakulteta za strojništvo, 2009.
- [4] Gebremedhin, A.: Optimal utilisation of heat demand in district heating system – A case study, *Renewable and Sustainable Energy Reviews* 30, 230–236 (2014).
- [5] Zhao, H., Holst, J., Arvastson, L.: Optimal operation of coproduction with storage, *Energy*, Vol. 23, No. 10, 1998, str. 859–866.
- [6] Gadd, H., Werner, S.: Daily heat load variations in Swedish district heating systems, *Applied Energy* 106, 47–55 (2013).
- [7] Potočnik, P., Soldo, B., Šimunović, G., Šarić, T., Jeromen, A., Govekar, E.: Comparison of static and adaptive models for short-term residential natural gas forecasting in Croatia, *Applied Energy* 129, 94–103 (2014).
- [8] Smrekar, J., Potočnik, P., Senegačnik, A.: Multi-step-ahead prediction of NOx emissions for a coal-based boiler, *Applied Energy* 106, 89–99 (2013).
- [9] Makridakis, S., C. Wheelwright, S., J. Hyndman, R.: *Forecasting: Methods and applications*; United States of America, 1998.
- [10] Lee, C. M., Ko, C. N.: Short-term load forecasting using lifting scheme and ARIMA models, *Expert System with Applications* 38, 5902–5911 (2011).
- [11] Hyndman, R. J., Koehler, A. B.: Another look at measures of forecast accuracy, *International Journal of Forecasting* 22, 679–688 (2006).
- [12] Bacher, P., Madsen, H., Aalborg Nielsen, H., Perers, B.: Short-term heat load forecasting for single family houses, *Energy and buildings* 65, 101–112 (2013).
- [13] Dotzauer, E.: Simple model for prediction of loads in district-heating systems, *Applied Energy* 73, 277–284 (2002).
- [14] Aalborg Nielsen, H., Madsen, H.: Modeling the heat consumption in district heating systems using a grey-box approach, *Energy and buildings* 38, 63–71 (2006).
- [15] Kwak, Y., Seo, D., Jang, C., Huh, J. H.: Feasibility study on a novel methodology for short-term real-time energy demand prediction using weather forecasting data, *Energy and Buildings* 57, 250–260 (2013).
- [16] Diebold, F. X., Mariano, R. S.: Comparing predictive accuracy, *Journal of Business & Economic Statistics* 13, 253–263 (1995).
- [17] Soldo, B., Potočnik, P., Šimunović, G., Šarić, T., Govekar, E.: Improving the residential natural gas consumption forecasting models by using solar radiation, *Energy and Buildings* 69, 498–506 (2014).

Research and development of empirical forecasting models for heat consumption in the district heating system

Abstract: Forecasting energy consumption in district heating networks is crucial for wise energy resource management and sustainable energy supply. We investigated which factors most affected the heat consumption and how the forecasting accuracy can be improved. We found that air temperature, past heat consumption and solar radiation had the greatest impact on heat consumption. Nine empirical models were developed for heat consumption forecasting in the district heating system. Random-walk model and temperature correlation model were applied as reference models, and then more complex regression model and several types of autoregressive models were developed. The quality of empirical models was estimated by cross validation procedure based on forecasting error which represented the mean absolute error normalized by the maximum transfer capacity of the district heating network. Forecasting results have been improved by introducing the population variables for the days of the week and additional variables describing the system growth and seasonal cycle. The best forecasting result was obtained by an empirical model from a group of autoregressive models (model SR), which was developed by a stepwise regression method.

Keywords: District heating, heat load forecasting, forecasting models, feature extraction, feature selection, regression model, autoregressive model