Springer Proceedings in Energy

Ahmad Vasel
David S.-K. Ting *Editors*

The Energy Mix for Sustaining Our Future

Selected Papers from Proceedings of Energy and Sustainability 2018

Springer

Springer Proceedings in Energy

More information about this series at http://www.springer.com/series/13370

Ahmad Vasel · David S.-K. Ting
Editors

The Energy Mix for Sustaining Our Future

Selected Papers from Proceedings of Energy and Sustainability 2018

Springer

Editors
Ahmad Vasel
Fluid Mechanics Research Laboratory,
Mechanical Engineering Department
Tennessee Tech University
Cookeville, TN, USA

David S.-K. Ting
Turbulence and Energy Lab
University of Windsor
Windsor, ON, Canada

ISSN 2352-2534 ISSN 2352-2542 (electronic)
Springer Proceedings in Energy
ISBN 978-3-030-00104-9 ISBN 978-3-030-00105-6 (eBook)
https://doi.org/10.1007/978-3-030-00105-6

Library of Congress Control Number: 2018954490

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The pragmatic way forward to supply energy for a future quality living has to be threaded on a fusion path, where conventional energy is unceasingly refined as major renewable energy is made more financially sustainable. This proceeding contains a highlight of preeminent advances of energy, including a touch on the associated politics.

A general and largely unbiased (unpolitically corrected) disclosure on the policy, politics, and practice of energy and sustainability is ushered by Reader in Chapter "Energy and Sustainability: Policy, Politics and Practice". This is immediately followed by a detailed modeling of proton exchange membrane fuel cell system by Yang et al. Along the more established engine conversion systems, the flow field around a spark plug is critical as we push for even cleaner internal combustion engine, while simultaneously boosting fuel efficiency; see Chapter "Preliminary Simulation Study of Flow Field Around a Spark Plug Under Ambient and Engine Conditions" by Sandhu et al. The higher thermal efficiency compression ignition engine is possibly most vital for heavier machineries such as the transport truck. Yang et al. highlight one such latest progress in Chapter "Combustion Characteristics and Emissions of Direct Injection Neat n-Butanol in a Compression Ignition Engine".

Moving forward, we must make more room for renewables such as hydro-turbines. Adhikari and Wood propose a niche angle to overcome the age-old challenge of water turbines, i.e., cavitation. Solar must soar, and photovoltaic thermal (PVT) system is "the best of both worlds." Burhan et al. beautifully expound on the triple-pronged approach, marrying hybrid concentrated photovoltaic system with hydrogen production in Chapter "Compact CPV—Sustainable Approach for Efficient Solar Energy Capture with Hybrid Concentrated Photovoltaic Thermal (CPVT) System and Hydrogen Production". What about furthering current photovoltaic system? Simple and effective passive cooling via the use of a square wire at the optimal attack angle can be cost-effectively implemented, as discussed by Yang et al. in Chapter "The Influence of Square Wire Attack Angle on the Heat Convection from a Surrogate PV Panel".

Sustaining a green environment must include sustainable and healthy food. For temperate places, in particular, versatile greenhouses with easily adjustable sides called high tunnel greenhouses can be furthered to extend the growing season. Ghose and Lubtiz detail the integrated multiple parameter tuning to furnish the most desirable microclimate in Chapter "Predicting the Interior Conditions in a High Tunnel Greenhouse".

Energy conservation and recovery are crucial to the sustainability of the ecosystem which includes the homo sapiens. What is a better way to educate future generations of this hard truth than to execute a wastewater heat recovery system on campus? Thus, it is the case presented by Chow and Duquette in Chapter "Assessment of a Heat Pump-Based Wastewater Heat Recovery System for a Canadian University Campus".

The timely gathering of varied experts at the Energy and Sustainability 2018 Symposium and Industry Summit, Windsor, Ontario, Canada, ought to continue. Perkins provides the frameworks to continue the sociopolitical discussion for public policy and energy investment decisions in Chapter "Sustainability and Energy Services: A Framework for Discussion". This concluding chapter also serves as a transition and invite to Energy & Resources for Tomorrow 2019; see https://eart2019.ca/.

Ahmad Vasel
David S.-K. Ting

Contents

Energy and Sustainability: Policy, Politics and Practice

Graham T. Reader(✉)

Department of Mechanical, Automotive and Materials Engineering, University of Windsor, 401 Sunset Avenue, Windsor, ON, Canada
greader@uwindsor.ca

Abstract. There is no universally accepted definition of what 'Sustainable Energy' means but in general the concept is to achieve a supply of energy that is sustainable over long periods of time with manageable or no negative environmental impacts. Unfortunately, many of these phrases are vague and open to interpretation. For example, when referring to long periods of time, do we mean decades, centuries or even longer? To be sustainable does an energy source have to be renewable and renewable forever? If the latter then no such source exists even the sun has consumed 50% of its hydrogen fuel. So, presumably what constitutes renewable energy depends upon the time frame over which the energy is to be used and the rate of use? What do we mean by manageable environmental impacts? Does this mean we are prepared to accept some negative impacts—whatever 'negative' means?

National concerns about energy have, until recently, been associated with the supply, security and cost of energy. Only as the global human population has dramatically increased, especially since the 1950s, have the environmental impacts of energy use become dominant concerns. The energy sources which have come under the most intense scrutiny are fossil fuels which account for almost 89% of global energy consumption and whose use contributes to, measurable increases in atmospheric carbon dioxide emissions, which are causing the retention of too much solar energy resulting in detrimental 'global warming', increases in the number of disruptive weather occurrences and rises in sea levels. Together these effects could result in over half the earth's population having to move away from coastal regions, whilst longer growing periods in other global regions would result in enhanced agricultural food production. Can all these effects be negated by a relatively immediate cessation of fossil fuel usage? How quickly could 89% of global energy consumption be eliminated or replaced and in a sustainable and affordable manner?

At least for the next few million years, solar and wind energy will provide obvious sources of renewable and sustainable energy. However, in terms of the electrical power generation, these sources are ill-suited for baseload generation, unless large-scale affordable energy storage systems can be developed. In the meantime, hydroelectric, nuclear and especially fossil fuel power generation will still be needed. Realistically, the latter is a finite energy source and for example, since the 1970s, forecasts have been made envisaging that oil reserves will be totally exhausted within a period of 10 to 20 years from the time of the particular predictions. Nevertheless, fossil fuels usage has not gone away—will they ever?

In this presentation, the historic and modern pathways of energy use are discussed along with the accompanying environmental concerns and 'no-harm'

A. Vasel and D. S.-K. Ting (Eds.): EAS 2018, SPE, pp. 1–17, 2019.
https://doi.org/10.1007/978-3-030-00105-6_1

energy sustainability which have become more focused especially in the last half-century. However, in the face of polarized political, economic and societal opinions, will it be possible to achieve an agreed global plan for universal sustainable energy implementation in the near future? Can science and technology alone provide successful solutions to our energy and environmental dilemmas?

1 Introduction

According to sustainable energy values or beliefs, the Earth needs an energy source or sources to support the needs of the human population now and forever in the future, and whose use and acquisition causes no harm to the environment whilst concomitantly not diminishing the source or sources for future generations. How can this scenario be achieved? In the first instance, it would be useful to know what energy sources are used by the global human population now and in the past since if these energy sources do not meet the criteria as outlined then they will need to be replaced. If the replacements are more expensive than current sources, then it seems unlikely that our species will readily adopt them without a governance policy lawfully instituted by a democratically elected government or by edict from an absolute ruler. The former may prove more difficult than the latter especially when there is a strong culture of bipartisanship and the omnipresent requirements of the 'loyal opposition' to opposite, futilely, just about every government decision or proposal. Moreover, if the proposal disaffects too many voters a government may decide, they wish to remain in power, not to pursue the original proposal to its fullest extent and vice versa, i.e. a 'lobbied' proposal may gain traction for the same reason.

A 'lobbied' proposal may be the result of vested interests of varying complexions ranging from profit-based motives, to political and philosophical ideologies, to increasing individual and group government research grants. In recent years, we have been strongly encouraged to believe or at least trust 'the science', exactly what science we should be trusting remains problematic. Nevertheless, accurate measurements of global atmospheric average temperatures have indicated that they have increased in 16 of the last 17 years and if the rate of increase continues to increase without some actions being taken, the consequences could be dire not only for humanity but all living things. Even so predicting the future is usually precarious, for example, just before the events occurred leading scientists and philosophers predicted that flight by machines was 'unpractical … if not utterly impossible' and that scientific lecturers 'demonstrated to satisfied audiences that a ship could not cross the Atlantic by the power of steam' [1, 2]. We also have to contend with the 'Turnbulls' encountered in Trollope's *Phineas Finn,* 'Mr. Turnbull had predicted evil consequences …and was doing the best in his power to bring about the verification of his own prophecies' [3]. If then sustainable energy is to be taken seriously and capture the attention of all, we need to ensure that the actual issues regarding energy—the needs and uses, the associated environmental and societal impacts of non-sustainable and non-renewable energy usage and how continued human population growth affects global energy and the environment—are well and understood and appreciated. As David Suzuki stated four decades ago, 'If we

are to have any hope of controlling the elements that will transform our lives, an understanding of science is imperative' [4]. Perhaps, one of the first actions is to agree exactly what we are talking about since we have gone through mantras starting in the 1950s with Hubbert's peak oil theory followed by oil crises in the 1970s with lead to claims of rapid oil reserves depletion [5], followed by global cooling, then global warming and then climate change and weather disruptions. The latter is often confused in the media where climate change is used as the reason for a weather event, especially an extreme weather event. Whereas climate is concerned with very long-term averages, weather is about what is happening now and in both cases geographical location appears to be a crucial factor [6]. Similarly, the Climatologists (Climate Changes) and Meteorologists (Weather Forecasting) need to use 'models' to predict future climate and weather, although the term 'model' is rarely used in the popular and non-technical media, perhaps to avoid the dreaded populist showstopper of 'theory' entering the discussions! Maybe we should have adopted the ancient Greek term, 'Paradigm' and 'Paradigm shift', as promoted in the seminal book 1960s book by physicist Kuhn [7], when discussing energy and the environment in the same way as the terms have become commonplace when discussing business, social movements, education, science and so on?

There are three dominant fossil fuels, coal, natural gas and oil, and each provides roughly equal amounts of the global energy consumption, Fig. 1. So why the apparent focus on oil and diminishing global oil reserves and the media's minute by minute tracking of spot oil prices and not coal and only rarely natural gas?

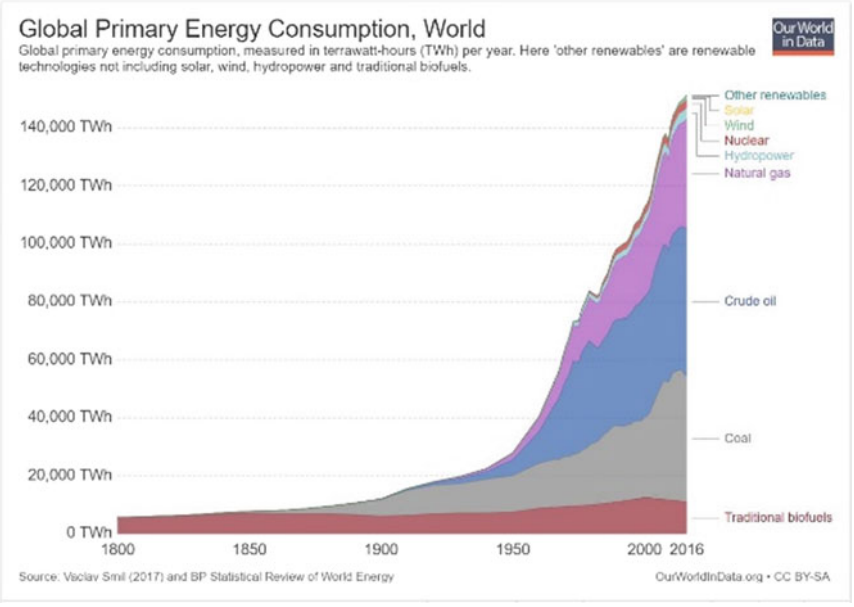


Fig. 1. Global primary energy consumption 1800–2016, OurWorldData.org.

The most probable reason is that oil is the fuel of choice for all forms of powered transportation—land, sea and air. Personal transportation is very important to almost all humans especially in developed countries. In a 2015 survey [8], automobile vehicle ownership was over 80% in South Korea (83%), Japan (81%), America (88%) and topping the list Italy (89%). China has become the world's largest new car market with ownership of cars and motorcycles increasing twentyfold in the first decade of this millennium. It has been estimated that global ownership of personal on-road vehicles is over 1 billion with some 250 million registered in the USA alone [8]. The global focus on oil is then perhaps not surprising. What is surprising is the general misunderstanding of what is meant by oil reserves as typified by Jimmy Carter's Presidential pronouncement in 1977 that, 'If it were possible to keep it {*global oil consumption*} rising …… as it has in the past, we could use all the proven reserves of oil in the entire world by the end of the next decade' [5]. As we shall see later in this paper, 'proven reserves' is an economic definition, which is also influenced by technology, and can be applied to all natural resources not just oil. The amounts of a particular natural resource and the proven reserves of that resource are quite different. The longevity of possible fossil fuel usage could then be significantly greater than usually forecast. Most certainly, fossil fuel is arguably a finite natural resource and therefore not in harmony with the designation of a sustainable energy source. So why the rush to replace fossil fuels especially in personal vehicle powertrains? It is apparent that the main thrust is the concern of the environmental impacts of oil-fuelled vehicles which produce unacceptable levels of greenhouse gases, criteria pollutants and carbon dioxide which in turn significantly contribute to anthropogenic effects on climate change [9] which the majority of climate scientists consider to be adverse and damaging to the 'health' of our planet [10].

So it appears that most policymakers, to conciliate the predictions of the climatologists, must promptly if not abruptly institute a 'paradigm shift' away from using almost 90% of the current type of energy source usage and convince at least 1 billion owners to replace their vehicles or at least the powertrains. While it is unlikely that the timeframe will be to the predilection of the scientists and environmental activists, if policymakers are to be successful, then the development of sustainable and affordable energy sources to replace fossil fuel sources is paramount. But is sustainability as crucial to the public as affordable non-carbon alternatives? The most obvious sustainable energy source is solar power as it has been calculated that the sun will not physically destroy the earth for about 7.5 billion years but its heat will render the earth uninhabitable for humans in about a billion years (not forever a renewable and sustainable energy source but close). However, for a fifth of the global population, 1.5 billion, i.e. those living in South Asia, some researchers have predicted that unless 'serious' mitigation of climate change occurs there is a unique risk of the region becoming essentially non-survivable for humans by the end of this century [12]. However, if serious mitigation means stopping using fossil fuels, immediately the literature indicates that 'global warming' will continue for at least several decades and although global average temperatures will eventually stabilize, the level will be higher than in the past [13]. As solar power is still more expensive—mainly because of the initial infrastructure costs—it will be difficult to persuade users to accept revolutionary changes which for most will have no obvious effect in their lifetimes.

Wind energy is also a candidate as a sustainable and renewable energy source as it is largely dependent on the effects of solar radiation. However, like solar power, it is still too expensive in comparison with its fossil fuel competitors and not all are convinced that it is a no-harm environmental technology. Nevertheless, the US Energy Information Administration (EIA) has indicated that wind power is expected to exceed hydropower in terms of electrical generation by 2020 [14], Fig. 2. Solar and wind energy sources are realistic as sustainable replacements for fossil fuel sources but if wholly implemented would require all modes of transportation to be electric. This would be easier to accomplish for land transportation but an increasing challenge for marine transportation—which carries at least 80% of the world's trade [15]—and air travel. Approximately, 30% of seaborne trade is the transportation of oil and gas so with the paradigm shift this would be no longer needed but that would leave over 7 billion tonnes of cargo to be transported [15]. At the moment, marine transportation accounts for 3–4% of global carbon dioxide emissions. Is it realistic to expect all shipping to be electric and/or wind powered immediately or at least in the next decade? Whatever edicts are made by national and international agencies, and they must be unanimous to have the desired impact, will such a very minor reduction of carbon dioxide ameliorate climate change? Of course, eventually, marine transportation will have to seek alternative non-carbon energy sources but until electrochemical batteries can be developed that can power a 100,000 tonne or more ship and its auxiliaries (e.g. refrigerators) for an average sea journey of 10–15 days, it seems unlikely that fossil-fuelled shipping will be replaced any time soon.

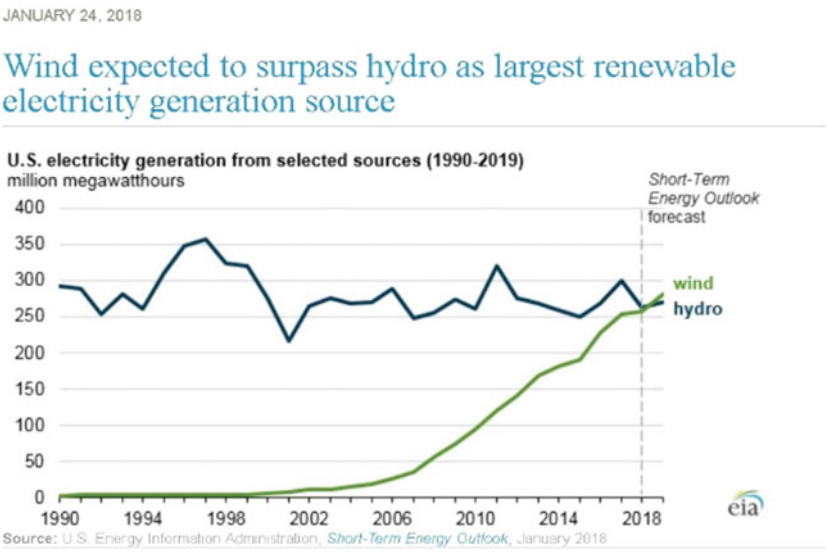


Fig. 2. Wind versus hydropower in US electricity generation, USEIA January 2018.

2 The Rise and Fall (?) of Fossil Fuel Usage

Fossil fuel usage—coal—is usually associated with the British Industrial Revolution of the eighteenth century onwards. The 'success' of this revolution was founded on access to cheap coal and high labour wages. It appears that all the current and future problems regarding global climate change are the result of this and similar industrial revolutions and that they are no positives to attach to these occurrences. The increases in national and individual wealth, improvements in quality of life, education and medicine, human population growth and life expectancy, Fig. 3, are no excuse, it seems, for causing Climate Change especially for future generations.

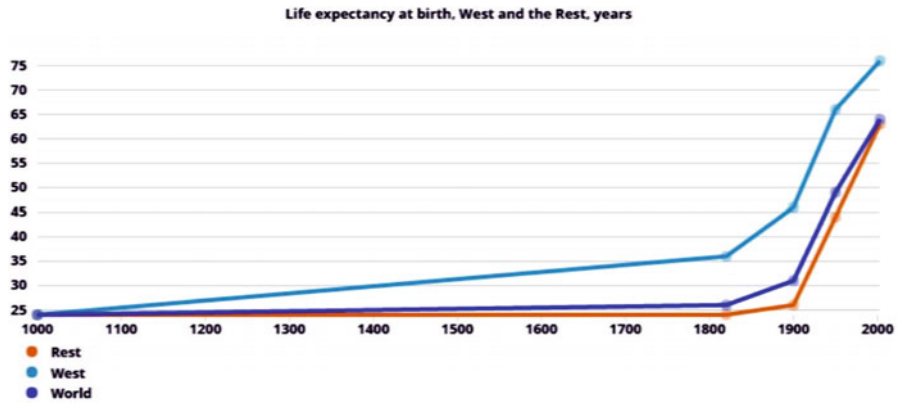


Fig. 3. Life expectancy at birth 1000–2000 CE, [16].

An argument could be made that many of those making the negative observations about fossil fuels usage would probably never have been born or maybe they would have been dead by the age of 27 to 35, and most likely in their lifetimes they would not have enjoyed access to post-secondary education or indeed any form of education, if such fuels had not been exploited. While there is some substance to these views, nevertheless the potential economic and social costs of climate change could be colossal from the twenty-second century onwards. Regardless of the cause and effects, returning to human muscle power (as often as not as a slave), animal muscle power, water power and wood power may appeal to some, it is unlikely that it would be universally endorsed.

As the modern industrial revolution started in Britain maybe some insight into why they chose—evidently disastrously—to burn 'sea' coal rather than wood [17]. We can trace this choice back to some localized locations in Roman Britain but especially in the thirteenth and early fourteenth centuries coal use became increasingly common [17–19]. At that time, the population in Britain was twice as high as it would be at the start of the industrial revolution. To feed this, growing population required more land for agriculture—also blamed by some for causing climate change—and as forestry management was rudimentary in many northern areas of England, to say the least,

wood for heating and cooking became scarce and increasingly expensive. The 'better off' landowners could afford to pay the higher prices—but the majority of the population—could not. Indeed, the rich largely abhorred the burning of 'toxic' coal because of the smoke and smell and has early as 1257 Queen Eleanor was forced to leave Nottingham Castle probably because of 'air pollution' according to Te Brake [17]. Eventually, because of the ready availability of sea coal and its lower cost, the wealthy allowed the medieval lime kiln operators and blacksmiths to use it along with their 'serfs'—the poor Villeins—for domestic heating and cooking. The latter use helped reduce the increasing mortality rate among the serfs. Why *sea* coal? Although the general consensus among historians was that coal was referred to as sea coal because it was transported by sea from the Northern UK coasts to places like London, there is evidence to suggest that the origins of the title came about because coal was found near or just under the sea [17].

However, in 1273, the use of coal was prohibited in London 'as it was prejudicial to health' but by 1306 a Royal proclamation only prohibited certain types of coal. The legislation was to address the complaints about air quality [20]. Nevertheless, in most rural areas and at times in the cities, coal burning increased. So why by the mid-to-late fourteenth century did coal use largely disappear until the seventeenth century? The answer is the bubonic plague—the Black Death—which swept across Europe and Britain in successive waves causing the demise of upwards of a half the population with one of the ramifications being that there were no longer any wood shortages [21]. Only by the start of the eighteenth century did the population reach the levels of the early fourteenth century and by then wood was also needed for shipbuilding as Britain was fast become a global sea power [22]. The technology of steam power was also rapidly advancing and perhaps ironically was put to use to drain quarries and mines to enable deeper coal seams to be worked. Eventually, the start of the modern manufacturing factory and the advent of steam-driven rail transportation consolidated coal as the energy source of choice for all.

Industrialization also led to increased urbanization, a typical example being the British state capital London. At the start of the nineteenth century, the city had a population of 1 million which doubled by mid-century and reached 6.5 million by the end of the century, 20% of the country's total population being squeezed into a relatively small space. It was an unhealthy place to live especially for the poor- and low-income earners because of polluted water and air, the 'better off' moved to the suburbs. The relationship between pollution and disease, essentially 'germ theory', was not believed until the end of the century approached. Clean water became a priority and a network of public parks was established to improve air quality but one of the biggest 'killers' was lung disease [23]. Throughout the nineteenth century, many British cities experienced 'Smog'—pea soupers—and although the number of incidents declined in the twentieth century, the laws were rarely enforced and the general consensus was that the smoke from coal burning was not a big issue. Then, in December 1952, after a period of cold weather, an anticyclone settled over London which caused the wind to drop and the air to become damp, forming a thick fog. This turned into smog as sooty particles became trapped. The anticyclone and smog lasted for 5 days—causing an estimated 4,000 to 6,000 deaths and as many as 100,000 cases of respiratory illness were reported. This lead to the UK's Clean Air Act in 1956 and the Tall Chimney

legislation in 1968, requiring the establishment of 'smokeless' zones, the production of cleaner (reduced Sulphur content) coal and clean coal technologies such as smokeless coal fuels and higher chimneys.

Almost seven centuries after, coal usage legislation was first introduced Government started to take the air pollution issues seriously and enforce and implement the laws. This was not as difficult as might be imagined since both the coal and electrical generation industries were nationalized, i.e. 'Government' owned. Today, other than a few opencast sites, there is no coal mining industry in the UK and electrical generation has been privatized. However, starting in the mid-nineteenth century, a new 'alternative' fuel which at the time was somewhat expensive came to market—petroleum (oil). With the advent of the use of the Internal Combustion engine for vehicular propulsion in the early twentieth century and cheaper oil products, gasoline, petrol and diesel alternatives to coal were now readily available albeit they were also fossil fuels. In the same way, as cheap coal had driven Britain's economy to the premier league of global economies, oil would do the same for the USA. Gradually, oil became the energy source of choice for many applications, for example, the decision made by the Royal Navy to convert from coal to oil—lead by the future Prime Minister Winston S Churchill—lead to other navies doing the same although most had no known national oil resources. In due course, oil exports would be used by the USA for geopolitical purposes as part of economic sanctions against the Japanese after their armed invasion of the Chinese mainland in 1937. The Asia-Pacific war broke out 4 years later culminating in the use of thermonuclear weapons.

Like coal, the use of oil and natural gas has a far longer history than many imagine. The Mesopotamians used oil sources, and the famous Greeks, Aristotle and Herodotus wrote about such deposits almost three millennia ago and in more recent times a seventeenth-century French missionary wrote about Seneca First Nation using oil burning in their religious ceremonies [24]. In almost every case up to the eighteenth and nineteenth centuries, oil and natural gas were used solely for lighting purposes and very occasionally as lubricants for axels of horse-drawn vehicles. However, chemists were able to separate paraffin (kerosene) and by the mid-century chemical refining—especially in the USA—was enabling petroleum fuel products to be produced which could be used in dynamic heat converters.

The USA and, to a lesser but not insignificant extent, Venezuela became the main oil exporting countries with 75% of the world's coming from the US well into the twentieth century but by the 1980s Saudi Arabia was producing more oil than the US and by 2015 Russia was producing more oil than Saudi Arabia, Fig. 4 [25]. Today, around 100 different countries have oil production of some sort. The use of a type of third fossil fuel—natural gas—rapidly escalated following the invention of the 'Bunsen Burner' which enabled its use for many applications other than lighting, although the Chinese had used natural gas 2,500 years ago in their desalination plants. Initially, in the first flush of the industrial revolution, natural gas was manufactured from coal on a commercial basis in 1785 by Britain to be used for residential and street lighting. Four decades later, industrial extraction of natural gas—usually found in the vicinity of other fossil fuels such as oil and coal—began in the USA. Today, like oil extraction, natural gas production involves many countries with North America, Europe and Eurasia still the major sources, Fig. 5 [26].

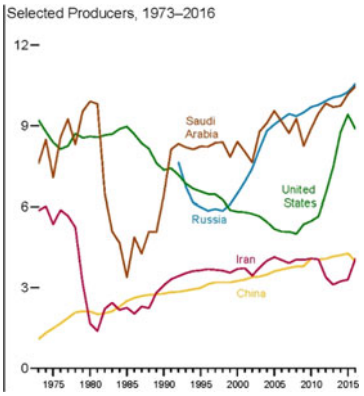


Fig. 4. Major oil producers [25]

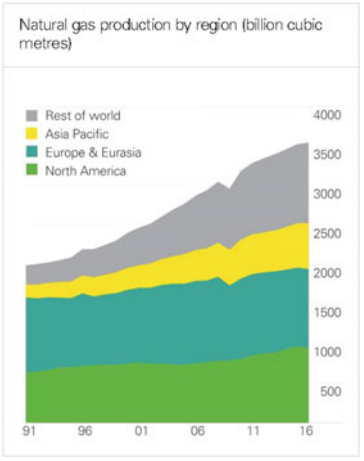


Fig. 5. Natural gas production (1991–2016) [26].

So it would appear that the production of fossil fuels, in particular, liquid oil and natural gas, is increasing not decreasing. Does this mean, for example, that the global 'proven' reserves of these fuels will soon be depleted? The epithet, 'proven', is often dropped in discussions about fossil fuels especially in the popular media. Maybe this is intentional or maybe the concept of 'proven' in relation to fossil fuels is misunderstood. The use of the term 'reserves' can give the impression that their stated amounts provide a complete picture, a total, of all there is left. This is not the case. Proven reserved are defined as, 'fossil fuels that are technically and economically recoverable at current prices' [27]. If for example oil prices increase, then it could be reported that proven reserves have increased, whereas lower prices could result in decreases in estimated reserves. The oil is still in the ground but may or may not be included in the amounts of proven reserves because of economically rather than geological reasons. Thus, until

relatively recently, Canadian oil sands and oil and gas from shale deposits were not included in the proven reserve calculations. The situation is further complicated by the fact that 'shale oil' and 'oil shale' are different entities and while shale oil—which refers to trapped hydrocarbons and accessed by 'fracking'—is starting to be included in proven reserves calculations, and oil shale is obtained from rocks containing kerogen—a precursor to oil. The technologies for producing oil from mined rock are still too expensive although oil shale rocks are plausibly an enormous source of oil and gas [28]. Over 90% of shale oil and gas resources exist outside the USA, and shale gas represents a third of the total global gas resources.

Economics aside, geologists also calculate 'resources', i.e. known fossil fuel deposits that exist but whose extraction is not economical at current prices. These resources also include methane hydrate—a mixture of natural gas and water usually found beneath sea floors. The difference between the amounts of 'resources' and 'proven reserves' is a factor of 2.8 to 4 for oil, 4 to 58 for natural gas and 14 to 23.5 for coal [29]. If the resources become proven reserves, then at the current rate of usage, coal could last for over 2,000 years, gas for over a 1,000 years and oil into at least the twenty-first century. While these are unlikely scenarios, it has to be remembered that new deposits and sources of fossil fuels are being found on a regular basis which partly explains why, for example, the calculated timeframe for the complete depletion of oil has not changed since the 1970s despite increased consumption. The ratio of consumption to proven reserves has remained sensibly constant as depicted in Fig. 6.

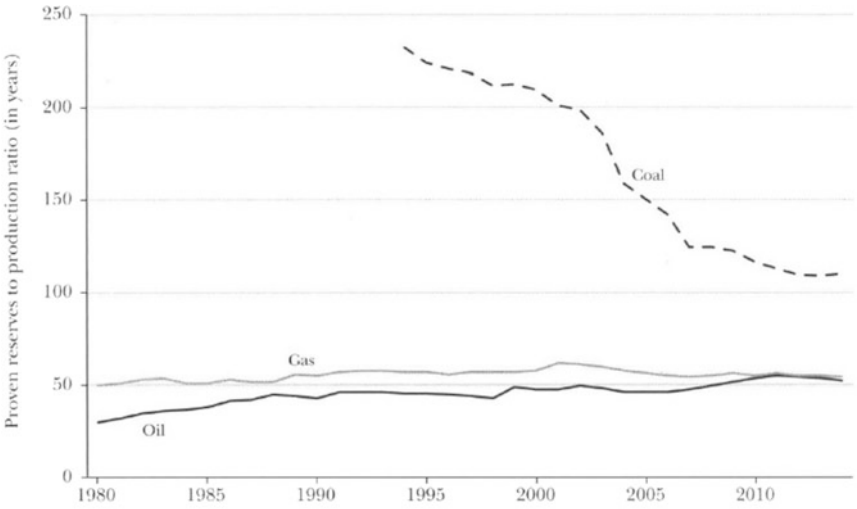


Fig. 6. Ratio of proven reserves to consumption [5].

Even if all the fossil fuel resources became 'proven' reserves, their longevity of up to a millennium or two would not qualify as defined 'sustainable energy'. Nevertheless, the diminishing reserves and resources arguments, once the facts are objectively

delineated, may not be as convincing—at least to the public and future government administrations in a number of countries—including the top three oil producers. Consequently, if there is to be a paradigm shift away from fossil fuel usage, then the potential costs of the non-mitigation of anthropogenic climate change will be enormous and outweigh any of the benefits accrued in the past, the present and the future of using fossil fuels. The lack of longevity contentions is unlikely to be as persuasive as clearly validated cost and health issues. The health hazards associated with polluted water and air eventually gained general attention but even then it required enforced government action to address the problems, which in many parts of the world still exist. Prior to the greenhouse gas/climate change issues becoming recognized significant issues, Governments in North America, for example, had defined 'Criteria Pollutants' (USA) and 'Air Contaminants' (Canada). The two national lists are not identical and the same is true for the 'lists' of another nations; however, they were and are a number of commonalities on all lists, especially with regard to fossil fuel emissions from electrical generation and transportation, as shown in the next section, Table 1.

Table 1. Greenhouse gases, US pollutants and Canadian contaminants as specified by national regulations.

Greenhouse gases	US-EPA criteria pollutants	Canadian air contaminants
Water Vapour (H_2O)		
Carbon Dioxide	Carbon Monoxide	Carbon Monoxide
	Particulate Pollution	Particulate Matter (PM)
Nitrous Oxide (N_2O)	Nitrogen Dioxide	Nitrogen Oxides
Ozone (O_3)	Ozone	Ground Level Ozone
	Sulphur Dioxide	Sulphur Oxides
	Lead (Pb)	Ammonia
		Volatile Organic Compounds (VOC)
Methane		
Chlorofluorocarbons and Hydrofluorocarbons		

3 Pollutants, Contaminants and Responses

The actual agreed harmful and greenhouse gases emissions are given in Table 1. Sustainable energy sources must, by definition, produce no such emissions or at the very least only produce them in manageable quantities, presumably to negate any future adverse effects on anthropogenic climate change and eventually reverse the impacts of fossil fuel usage.

It has to be noted that carbon dioxide is not on the pollutants and contaminants lists. Indeed, when the various clean air acts came into force in the last quarter of the twentieth century, the main concerns were PM emissions from coal-fired power stations and diesel engines and nitrogen oxides—often collectively referred to as NOx—from a

range of devices using fossil fuels and in particular gasoline-powered vehicles. Despite the British efforts, 'smog' was still a problem especially in large cities and urban areas in general. Thus, in North America, regulations were imposed whereby there would be only an allowable amount of tailpipe emissions of NOx. The Automotive engine manufacturers responded quickly by introducing exhaust gas recirculation (EGR) with the aim of lower in-cylinder temperatures to levels which chemically precluded the formation of NOx. This was partially successful but came at the cost of reducing fuel efficiency (e.g. miles per gallon, litres per 100 km and so on). The US Congress responded by enacting the requirement for the Corporate Average Fuel Economy (CAFE) to be enforced by 1978 onwards for cars and light trucks. The purpose of CAFE was to reduce energy consumption in view of the almost imminent demise of oil reserves and to ensure energy security for the USA which was starting to consume more oil than it was producing. The overall strategy was to aggressively increase the CAFE requirements by legislation.

The vagaries of chemical reactions and chemical thermodynamics (especially dissociation) meant that fuel economy could not be increased at the same time as NOx was to be reduced and that carbon monoxide and unburned fuel would always be produced regardless of engine design. This lead to the understanding that tailpipe emissions would need to be different from actual engine emissions so began the age of 'after-treatment'. Until this time, the tailpipe (or exhaust) pipe which carried away the engine out gaseous emissions would only involve a muffler or silencer to reduce the noise. The solution was to pass the engine out emissions through a catalytic converter containing chemicals which would react with the unburned hydrocarbons and carbon monoxide (two-way converter) and transform these unwanted emissions to carbon dioxide (then considered harmless) and water vapour. Eventually, three-way catalytic converters were developed to also deal additionally with NOx and in more recent years four-way converters have become available which also remove particulates from gasoline-powered vehicles [30]. This has become necessary as gasoline engines now use fuel injection rather than carburetted pre-mixed fuel–air mixtures. Such fuel delivery systems long used in diesel-powered engines inevitably produce particulate matter [31]. The success of the technologies to reduce harmful emissions is illustrated in Fig. 7 [32].

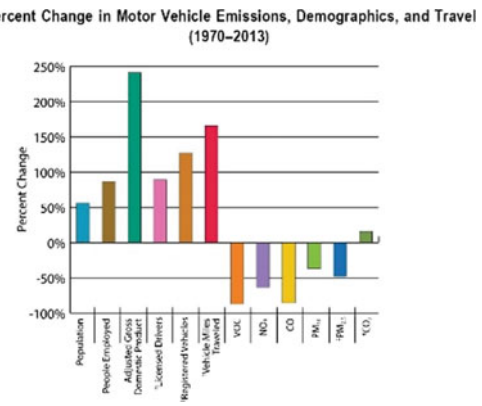


Fig. 7. % change if emissions, demographics and travel in the USA 1970–2013 [32].

Overall, the contributions of fossil-fuelled transportation to the global level of pollutants and contaminants is minimal especially for VOC, PM and NOx, whereas agricultural activities and fossil-fuelled electrical power generation and industrial processes are major polluters.

The systems developed to reduce the pollutants and contaminants from fossil-fuelled engines have achieved this by converting harmful emissions to water vapour, nitrogen and carbon dioxide. However, with the latter being identified as the major culprit in global warming and climate change the very success of the tremendous engineering feats achieved in the development of after-treatment systems—including the more recent use of Selective Catalytic Reduction (SCR) to further reduce NOx—is becoming their Achilles' heel. Solutions to this problem could include carbon dioxide sequestration but the practicalities of achieving this in personal transportation vehicles are probably insurmountable although sequestration in other fossil fuel applications has proved possible. But what to do with the sequestered carbon dioxide? One of the more intriguing ideas is to use solar power to convert the carbon oxide and emitted water vapour into gaseous fuels such as methane or synthetic liquid fuels [33]. This could be described as a type of mitigation approach and such approaches and are at least worthy of further investigation if the global average temperature is to be kept within the targeted maximum 2 °C rise [34] throughout the whole of the twenty-first century during the potential transition away from fossil fuels.

Although the efforts made by the majority of vehicular IC engine designers and manufacturers have been nothing short of miraculous—in engineering terms—in addressing pollutant, contaminant and fuel economy targets, which is why they have been used as an exemplar in this paper of what can be achieved, fossil fuel usage in this sector is but a part of overall picture [35]. If the consensus is that the climate change problem is wholly about carbon dioxide emissions from human activity, then the

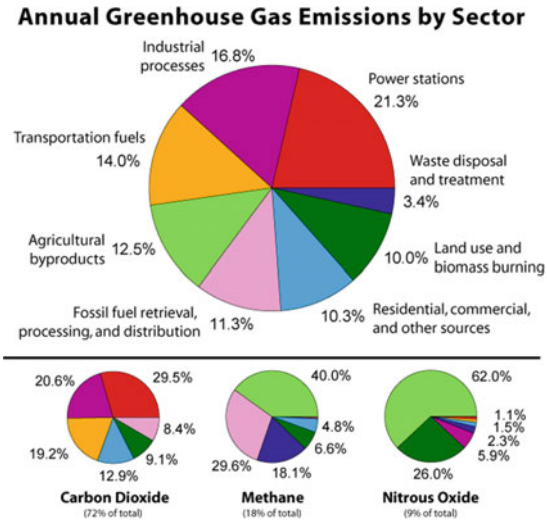


Fig. 8. Greenhouse emissions 2016 [35]

following figure, Fig. 8, depicts why the paradigm shift away from the current use of fossil fuels is of paramount importance.

4 Final Observations and Concluding Remarks

In a short paper such as this, it has only been possible to scratch the surface of the vast topic of sustainable energy sources. Yet, hopefully, it has been shown that no source meets the utopian definition of sustainable energy but that solar power (and by association wind power) and its development provide the most obvious way ahead. There is still much work to be done especially with regard to thermal and other forms of energy storage. The complete replacement of the energy sources of our present generation stations, industrial processes and transportation system is a gargantuan challenge and in the author's opinion, it is unrealistic to suggest that this could be achieved in the next 10 years as some academic researchers have advocated. Apart from the energy storage issues and the actual cost of solar power and its main conversion system—the solar panel—needs to be further reduced or the cost of oil has to double or triple in price to make the economic arguments persuasive to all concerned. Although great strides in the cost reduction of solar panels have been made this century, they are still too expensive, Fig. 9 [36], unless government incentives are evoked. However, in the USA, the price of solar panels to the consumers has been increased by the imposition, by the present administration, of a 25% tariff on imported panels. A policy decision which may help the mainland USA industry presently only meets about half the demand. Eventually, they may be able to address the potential shortfall but it is likely the panels will be more expensive and the policy will delay the transition to solar power. Ironically, the major US suppliers manufacture most of its components in Mexico.

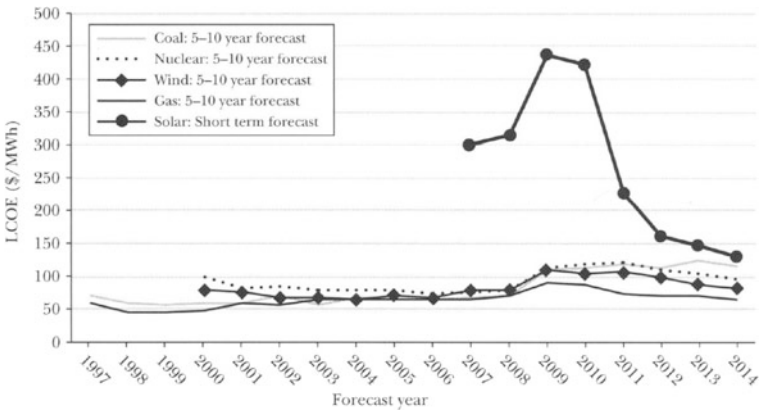


Fig. 9. Levelized Cost of Energy (LCOE) forecasts using USEIA data [36]

Solar panels and wind turbines do not last forever although 'warranties' for panels (if not all system components) can be obtained for at least 20 years and up to 40 years. They will then need to be replaced. However are end-of-life panels to be sent to landfill sites? As their use rapidly increases a lot more contaminated landfills sites will be needed unless the panels are recycled. Solar PV panels, however, contain heavy metals such as cadmium and lead as well as, in some cases, rare elements such as gallium and indium [37]. The panels are also constructed using a variety of different and disparate materials which will be a challenge to separate and recycle. Are recycling processes and facilities already in place?

The global automakers especially in Japan and North America have—usually because of legislation—developed sophisticated recycling processes and facilities which may be of some use to the solar power industry. These same automakers have reinvigorated the development of the electric car after earlier failures—caused as much by over-stringent regulations as technology—and while it seems that—at least in North America—the present electrical generation capacity coupled with smart grids could cope with a transition to electric vehicles only if such power generation is produced by non-carbon energy sources which will reduce carbon dioxide emissions. To replace all fossil-fuelled, and perhaps nuclear, generated electrical power is going to be a challenge and it maybe that natural gas-fired power stations would allow the cradle-to-grave carbon dioxide emissions of electrical vehicles to be lowered until solar, wind and hydro wholly replace fossil fuel power generation. Increases in nuclear power could also be instrumental in enabling the transition to electric vehicles but fears, largely unwarranted, of explosion catastrophizes and concerns about what to do with fission wastes will likely preclude nuclear playing a major role.

All current energy sources will become depleted in the future and it is unlikely that renewables will be able to be developed in time to meet all the global energy demands or all of the sustainability criteria. However, as far as we know solar energy, and by association wind energy, will be available for the next billion years or so and we should work towards its global use in a measured manner not by precipitous abandonment of current sources. Governmental policies will need to be pragmatic and unburdened with political ideologies. In practice, if there is no social buy-in on a global scale, then business as usual will prevail and anthropogenic damage to the climate will continue and increase. Of course, it has to be acknowledged that air pollution is not the sole preserve of fossil fuel use. Like the London smog of 1952, a 3-day weather inversion caused in 1948, the 'Donora' disaster in the town of Donora 39 km southeast of Pittsburgh in which 20 residents died and over 6,000 became sick. The disaster was caused by the emissions from the US Steel Zinc works being trapped by the inversion [38, 39]. According to the township, this incident was the 'birthplace of clean air' and instrumental in the passage of the Federal 'Clean Air Act'. In 2017, researchers looking for chemical markers of the disaster in local river sediments found none [38]. Unfortunately, the damage to the global heat shield caused by carbon dioxide emissions is going to last for far more than 70 years.

References

1. Newcomb, S.: Facts and Fallacies: A Book of Definitive Mistakes and Misguided Predictions. St. Martin's Press, New York (1981). (Ed. by, C. Morgan and D. Langford)
2. Murray, J.C.: Industrial Kingdom of God. University of Ottawa Press, Ottawa (1981). (Ed. by, L. Armourm and E. Torr)
3. Trollope, A.: Phineas Finn (Chapt. 25). The Oxford Dictionary of Quotations, 3rd edn, p. 553. Oxford University Press, Oxford (1979)
4. Suzuki, D.: CBC Radio Program Guide 16 December 1978. Colombo's New Canadian Quotations, p. 335. Hurtig Publishers, Edmonton (1987)
5. Covert, T., Greenstone, M., Knittel, C.R.: Will we ever stop using fossil fuels? J. Econ. Perspect. (Winter) 117–137 (2016)
6. https://www.accuweather.com/en/weather-news/weather-versus-climate-understanding-the-difference-of-the-two-terms/70003861. Accessed 31 May–9 June 2018
7. Kuhn, T.S.: The Structure of Scientific Revolutions. University of Chicago Press, Chicago (1962)
8. https://www.citylab.com/transportation/2015/04/global-car-motorcycle-and-bike-ownership-in-1-infographic/390777/. Accessed 10 June 2018
9. http://ipcc.ch/pdf/assessment-report/ar5/syr/AR5_SYR_FINAL_SPM.pdf
10. Periodically updated. https://climate.nasa.gov/causes/
11. Barras, C.: How long will life survive on planet earth? 23 March 2015. http://www.bbc.com/earth/story/20150323
12. Im, E.-S., Pal, J.S., Eltahir, E.A.B.: Deadly heat waves projected in the densely populated agricultural regions of South Asia. Sci. Adv. **3**(8) (2017). https://doi.org/10.1126/sciadv.1603322
13. Rood, R.B.: http://theconversation.com/what-would-happen-to-the-climate-if-we-stopped-emitting-greenhouse-gases-today-35011. Accessed 11 Dec 2014
14. https://www.eia.gov/todayinenergy/detail.php?id=34652. Accessed Feb 2018
15. http://unctad.org/en/Pages/Publications/Review-of-Maritime-Transport-(Series).aspx. Accessed June 2018
16. Lal, D.: Poverty and Progress, Realities and Myths about Global Poverty, 19 March 2014. https://doi.org/10.1111/j.1728-4457.2014.00663.x
17. Te Brake, W.H.: Air pollution and fuel crises in preindustrial London 1250–1650. Technol. Cult. **16**(3), 337–369 (1975)
18. Galloway, J.: The Coming of the Anthropocene, Unit 4 in Past Actions, Present Woes, Future Potential: The Higher Education Academy, November 2010. https://warwick.ac.uk/fac/cross_fac/heahistory/elibrary/internal/co_levene_pastactions_20100731
19. Fouquet, R., Pearson, P.J.G.: A thousand years of energy use in the United Kingdom. Energy J. **19**(4), p1–p41 (1998)
20. https://www.pollutionsolutions-online.com/news/air-clean-up/16/breaking_news/the_history_of_uk_air_pollution_looking_back_through_smog_tinted_glasses/32595. Accessed 29 Nov 2014
21. Galloway, J.: Op. Cit. [18]
22. Te Brake, W.H.: Op. Cit. [17]
23. Cook, B., Werner, A.: Breathing in London's History, 24 August 2017. https://www.museumoflondon.org.uk/discover/londons-past-air
24. Redwood, B.: Petroleum and its products, canton lecture part 1. J. Soc. Arts **11**, 805–814 (1886)

25. Monthly Energy Review: Figure 11.1a, May 2017. https://www.eia.gov/totalenergy/data/monthly/archive/00351705.pdf. Accessed 10 June 2018
26. https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy/natural-gas/natural-gas-production.html. Accessed 11 June 2018
27. Covert, T., et al.: Op. Cit. [5], p. 119
28. Beattie, A.: The difference between shale oil and oil shale, 28 May 2018. https://www.investopedia.com/articles/investing/080715/difference-between-shale-oil-and-oil-shale.asp
29. Covert, T., et al.: Op. Cit. [5], pp. 125–126, data compiled various government sources
30. https://www.basf.com/en/company/news-and-media/science-around-us/catalytic-converter.html. Accessed Dec 2017
31. Stone, R. (ed.): Introduction to Internal Combustion Engines, 4th edn, Palgrave, Basingstoke (2012). (Any modern engineering thermodynamics textbook/research monograph)
32. US Department of Transportation, Federal Highway Administration: Transportation Air Quality: Selected Facts and Figures, February 2016
33. Knight, H.: Green-machine-cars-could-run-on-sunlight-and-CO_2, 2 June 2010. https://www.newscientist.com/article/dn18993-green-machine-cars-could-run-on-sunlight-and-co2/
34. Rogej, J., et al.: Disentangling the effects of CO_2 and short-lived climate forcer mitigation. Proc. Natl. Acad. Sci. USA **111**(46), 16325–16330 (2014)
35. USEPA: Greenhouse Gas Emissions. https://www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data. Accessed 13 Apr 2017
36. Covert, T., et al.: Op. Cit. [5], p. 128
37. MacNeill, R.: Solar PV Module Recycling—Why It's Important, Renewable Energy World, 1 June 2016. https://www.renewableenergyworld.com/articles/2016/06/solar-pv-module-recycling-why-it-s-important.html
38. Nierenberg, A.: Everyone in Pittsburgh remembers the Donora disaster from 1948. Everyone, that is, except the Earth itself. Pittsburgh Post-Gazette, 22 August 2017. http://www.post-gazette.com/local/south/2017/08/22/donora-smog-1948-us-steel-zinc-works-markle-lake-pitt-forensic-analysis-environmental-disasters/stories/201707310050
39. https://www.aiche.org/chenected/2011/10/donora-death-fog-crisis-led-modern-air-pollution-laws

Modeling of Proton Exchange Membrane Fuel Cell System Considering Various Auxiliary Subsystems

Zirong Yang, Zhanrui Liu, Lichao Fan, Qing Du(✉), and Kui Jiao(✉)

State Key Laboratory of Engines,
Tianjin University, 135 Yaguan Road, Tianjin 300350, China
{duqing,kjiao}@tju.edu.cn

Abstract. In the study, a comprehensive proton exchange membrane fuel cell (PEMFC) system model is developed, including a two-dimensional transient multiphase stack model, a transient membrane humidifier model, air compressor model, and cooling model. The coupled water and heat transport processes, sophisticated water phase changes, gas/liquid transport in porous layers, and flow channels are taken into consideration in the integrated system model. Effects of gas purge duration in PEMFC stack and membrane humidifier on startup performance are investigated under subzero and normal temperatures. It is found that purge duration of membrane humidifier has little effect on output voltage when started from −10 °C and −5 °C since saturated vapor pressure is relatively small. Besides, the cold start duration is mainly determined by initial membrane water content in PEMFC stack. The upstream current density is usually higher because reactant gases are more abundant. To avoid sharp voltage drop during startup from 30 °C with large current density, long purge duration is not suggested for both PEMFC stack and membrane humidifier. The humidifier temperature is stabilized at about 42 °C as a result of exhausted gases heating and heat loss to environment when stack temperature is kept at 60 °C. The membrane water content in humidifier increases more rapidly when current density rises since more water vapor is generated and flows into humidifier; meanwhile, it results in higher humidifier temperature.

1 Introduction

Owing to its high power density, high electric efficiency, and zero emission, proton exchange membrane fuel cell (PEMFC) is widely recognized as one of the most promising energy sources in the future. To guarantee stable and efficient operation of PEMFC stack, various auxiliary subsystems are essential, including reactant gases supply subsystem, humidification subsystem, heat management subsystem, and control subsystem [1, 2]. The hydrogen/air supply subsystem provides sufficient reactant gases to stack which are subsequently consumed to generate electricity. The reactant gases are humidified by humidification subsystem before flowing into stack, maintaining good membrane hydration since dry membrane causes large ohmic voltage loss [3]. However, excessive humidification may result in water flooding, hindering occurrence of reactions, and transport of gas reactants. Heat management subsystem stabilizes

A. Vasel and D. S.-K. Ting (Eds.): EAS 2018, SPE, pp. 18–33, 2019.
https://doi.org/10.1007/978-3-030-00105-6_2

stack temperature in a reasonable range; besides, it controls temperature difference among individual cell. Control subsystem converts direct current from PEMFC stack to alternating current which can be used by the drive motor. Control subsystem is also capable of adjusting stack output power, protecting stack from severe fluctuations.

In the past decades, both experimental and modeling studies have been conducted to investigate the heat and mass transfer processes inside fuel cells. Much attention has been paid to effects of operating conditions, structure designs in membrane electrode design (MEA), and flow channel [4–8]. To cast more sights, visualization studies including scanning electron microscopy (SEM) [6, 9], electrochemical impedance spectroscopy (EIS) [10], infrared radiation imaging [11], and other high-tech methods [12] have been carried out. Besides, numerical methods such as Lattice Boltzmann method (LBM) [13], direct numerical simulation (DNS) [14], molecular dynamics (MD) [15], and computational fluid dynamics (CFD) [16, 17] have also been adopted. However, those models merely focus on fuel cell; they hardly take auxiliary equipment such as air compressor and membrane humidifier into consideration. Investigating the effects of operating conditions at a system level may result in different conclusions compared with conclusions drawn from individual PEMFC model or humidifier model.

Humidification subsystem has also been studied by numerous researchers, especially the operating conditions of membrane humidifier [18–20]. However, the humidifier models were developed based on either water-to-gas humidification or gas-to-gas humidification while both may happen in real PEMFC humidification subsystems, making it necessary to consider different humidification methods simultaneously. It is acknowledged that gas purge after the shutdown has great effects on fuel cell startup performance and lifetime [21, 22]. It was found that cold start survivability improved as purge duration increased. However, researchers also pointed out that membrane should not be too dry during maximum cold start mode in order to achieve fast startup [23]. In PEMFC system, both cold start and normal startup processes are coupled with membrane humidifier, which introduces additional water vapor to stack. The purging strategies of whole PEMFC system need further investigation, and effects of membrane humidifier on startup performance are also worth investigating since related studies are hard to find in literature.

In the study, a relatively comprehensive PEMFC system is developed, considering detailed mass and heat transport process inside fuel cell stack, transient behavior of membrane humidifier, air compressor, and cooling subsystem. First, the various submodels are compared with experimental data to validate its accuracy. Then, cold start performances under various startup temperatures and initial membrane water contents are investigated based on the system model. Besides, normal startup from 30 °C is also investigated. At last, effects of operating current density on membrane humidifier performance are investigated in order to obtain a better understanding.

2 Model Development

The schematic diagram of proton exchange membrane fuel cell (PEMFC) system is illustrated in Fig. 1, including PEMFC stack, gas supply subsystem, humidification subsystem, and heat management subsystem. Humidification is achieved by membrane

humidifier, and it is reported that cathode humidification is usually more important than anode humidification in previous studies [24]. The conservation equations of each subsystem are demonstrated in the following subsections.

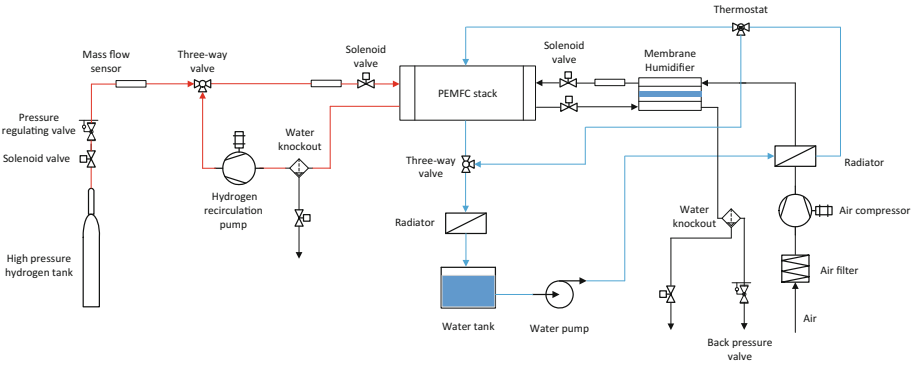


Fig. 1. Schematic diagram of proton exchange membrane fuel cell system considering various auxiliary subsystems.

2.1 PEMFC Stack Model

As shown in Fig. 2, the single cell consists of eleven layers, including bipolar plate (BP), gas channel (CH), gas diffusion layer (GDL), microporous layer (MPL), catalyst layer (CL), and proton exchange membrane (MEM). To enhance calculation efficiency, the membrane water content, liquid water volume fraction, gas concentration, and temperature are solved at the center of each layer with no further nodes divided. In order to depict gas distribution along flow channel, the single cell is further divided into several lumps, namely, the upstream and downstream. Gas temperature, concentration, and velocity in flow channel are transferred from the outlet of preceding node to the

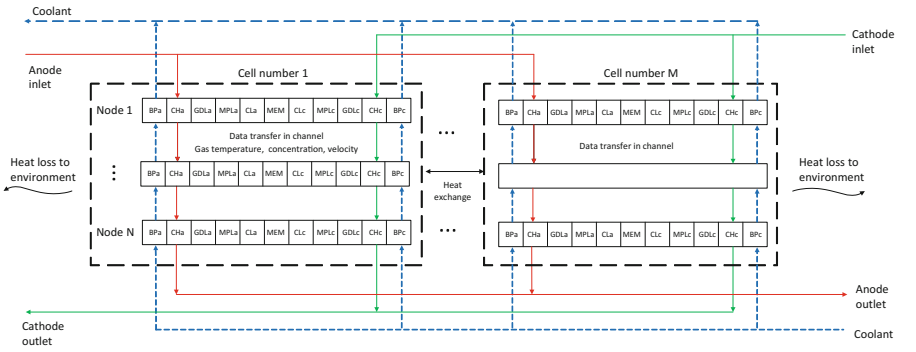


Fig. 2. Schematic diagram of the pseudo-two-dimensional transient PEMFC stack model.

Table 1. Cell properties and operating conditions.

Parameter	Value
Cell number	3
Nodes in flow direction	5
Effective area of single cell	100 cm^2
Channel length; width; depth; rib width	100; 1.0; 1.0; 1.0 mm
Thickness of membrane; CL; MPL; GDL	0.025; 0.01; 0.03; 0.2 mm
Stoichiometry ratio	ξ_a = 1.2; ξ_c = 2.0
Operating pressure	$p_{a,c}$ = 1.2 atm
Heat transfer coefficient between BP and surrounding	h_{surr} = 20 W m^{-2} K^{-1}
Heat transfer coefficient between BP and coolant	h_{cool} = 200 W m^{-2} K^{-1}

inlet of subsequent node, expanding the multiphase-layered model into a pseudo-two-dimensional model. The cell properties and operating conditions are briefly given in Table 1.

The membrane water content is calculated in membrane electrode assembly (MEA), which is updated directly at each time step.

$$\lambda_{\mathrm{CLa}}^{t}=\lambda_{\mathrm{CLa}}^{t-\Delta t}+\left(\frac{\left(\lambda_{\mathrm{MEM}}^{t-\Delta t}-\lambda_{\mathrm{CLa}}^{t-\Delta t}\right)D_{\mathrm{MEM\_CLa}}^{\lambda,\mathrm{eff}}}{\left(\frac{\delta_{\mathrm{CLa}}}{2}+\frac{\delta_{\mathrm{MEM}}}{2}\right)\delta_{\mathrm{CLa}}}+S_{\mathrm{mw}}\frac{EW}{\rho_{\mathrm{MEM}}}\right)\frac{\Delta t}{\omega_{\mathrm{CLa}}} \tag{1}$$

$$\lambda_{\mathrm{MEM}}^{t}=\lambda_{\mathrm{MEM}}^{t-\Delta t}+\left(\frac{\left(\lambda_{\mathrm{CLc}}^{t-\Delta t}-\lambda_{\mathrm{MEM}}^{t-\Delta t}\right)D_{\mathrm{MEM\_CLc}}^{\lambda,\mathrm{eff}}}{\left(\frac{\delta_{\mathrm{CLc}}}{2}+\frac{\delta_{\mathrm{MEM}}}{2}\right)\delta_{\mathrm{MEM}}}-\frac{\left(\lambda_{\mathrm{MEM}}^{t-\Delta t}-\lambda_{\mathrm{CLa}}^{t-\Delta t}\right)D_{\mathrm{MEM\_CLa}}^{\lambda,\mathrm{eff}}}{\left(\frac{\delta_{\mathrm{CLa}}}{2}+\frac{\delta_{\mathrm{MEM}}}{2}\right)\delta_{\mathrm{MEM}}}\right)\frac{\Delta t}{\omega_{\mathrm{MEM}}} \tag{2}$$

During cold start process, supercooled water freezing mechanism is adopted which has been explained in our previous studies [5].

The liquid pressure at the center of porous layers is calculated in a similar way.

$$p_{\mathrm{l,CL}}^{t}=p_{\mathrm{l,CL}}^{t-\Delta t}+\frac{dp_{\mathrm{l}}}{ds_{\mathrm{lq}}}\left(-\frac{\left(p_{\mathrm{l,CL}}^{t-\Delta t}-p_{\mathrm{l,MPL}}^{t-\Delta t}\right)D_{\mathrm{CL\_MPL}}^{\mathrm{lq,eff}}}{\left(\frac{\delta_{\mathrm{CL}}}{2}+\frac{\delta_{\mathrm{MPL}}}{2}\right)\delta_{\mathrm{CL}}}+S_{\mathrm{lq}}\right)\frac{\Delta t}{\rho_{\mathrm{lq}}\varepsilon_{\mathrm{CL}}} \tag{3}$$

$$p_{\mathrm{l,MPL}}^{t}=p_{\mathrm{l,MPL}}^{t-\Delta t}+\frac{dp_{\mathrm{l}}}{ds_{\mathrm{lq}}}\left(\frac{\left(p_{\mathrm{l,CL}}^{t-\Delta t}-p_{\mathrm{l,MPL}}^{t-\Delta t}\right)D_{\mathrm{CL\_MPL}}^{\mathrm{lq,eff}}}{\left(\frac{\delta_{\mathrm{CL}}}{2}+\frac{\delta_{\mathrm{MPL}}}{2}\right)\delta_{\mathrm{MPL}}}-\frac{\left(p_{\mathrm{l,MPL}}^{t-\Delta t}-p_{\mathrm{l,GDL}}^{t-\Delta t}\right)D_{\mathrm{MPL\_GDL}}^{\mathrm{lq,eff}}}{\left(\frac{\delta_{\mathrm{MPL}}}{2}+\frac{\delta_{\mathrm{GDL}}}{2}\right)\delta_{\mathrm{MPL}}}+S_{\mathrm{lq}}\right)\frac{\Delta t}{\rho_{\mathrm{lq}}\varepsilon_{\mathrm{MPL}}} \tag{4}$$

where r_{CH} is the ratio between channel area and cell area since gases cannot flow in or flow out through land area at the interface of GDL and channel.

Once liquid pressure is solved, capillary pressure and corresponding water volume fraction can also be solved because gas pressure is assumed to be constant for simplicity. The water volume fraction in flow channel is taken into consideration as well.

The capillary pressure in channel is calculated based on Leverett J-function, and it is used to calculate the diffusion term which determines water saturation in both GDL and channel at next time step to guarantee mass conservation.

The gas concentrations are updated at each time step, including water vapor, hydrogen, oxygen, and nitrogen.

$$c_{\mathrm{i,CL}}^{t} = c_{\mathrm{i,CL}}^{t-\Delta t} + \left(S_{\mathrm{i}} - \frac{\left(c_{\mathrm{i,CL}}^{t-\Delta t} - c_{\mathrm{i,MPL}}^{t-\Delta t} \right) D_{\mathrm{CL\_MPL}}^{\mathrm{i,eff}}}{\left(\frac{\delta_{\mathrm{CL}}}{2} + \frac{\delta_{\mathrm{MPL}}}{2} \right) \delta_{\mathrm{CL}}} \right) \frac{\Delta t}{\varepsilon_{\mathrm{CL}}} \tag{5}$$

$$c_{\mathrm{i,MPL}}^{t} = c_{\mathrm{i,MPL}}^{t-\Delta t} + \left(\frac{\left(c_{\mathrm{i,CL}}^{t-\Delta t} - c_{\mathrm{i,MPL}}^{t-\Delta t} \right) D_{\mathrm{CL\_MPL}}^{\mathrm{i,eff}}}{\left(\frac{\delta_{\mathrm{CL}}}{2} + \frac{\delta_{\mathrm{MPL}}}{2} \right) \delta_{\mathrm{MPL}}} - \frac{\left(c_{\mathrm{i,MPL}}^{t-\Delta t} - c_{\mathrm{i,GDL}}^{t-\Delta t} \right) D_{\mathrm{MPL\_GDL}}^{\mathrm{i,eff}}}{\left(\frac{\delta_{\mathrm{MPL}}}{2} + \frac{\delta_{\mathrm{GDL}}}{2} \right) \delta_{\mathrm{MPL}}} + S_{\mathrm{i}} \right) \frac{\Delta t}{\varepsilon_{\mathrm{MPL}}} \tag{6}$$

The temperatures at the center of all layers are calculated at each time step.

$$T_{\mathrm{BP}}^{t} = T_{\mathrm{BP}}^{t-\Delta t} + \left(\frac{\left(T_{\mathrm{CH}}^{t-\Delta t} - T_{\mathrm{BP}}^{t-\Delta t} \right) k_{\mathrm{BP\_CH}}^{\mathrm{eff}}}{\left(\frac{\delta_{\mathrm{BP}}}{2} + \frac{\delta_{\mathrm{CH}}}{2} \right) \delta_{\mathrm{BP}}} - \frac{h_{\mathrm{cool}} \left(T_{\mathrm{BP}}^{t-\Delta t} - T_{\mathrm{cool}} \right)}{\delta_{\mathrm{BP}}} + S_{\mathrm{T}} \right) \frac{\Delta t}{\rho c_{\mathrm{p}}} \tag{7}$$

$$T_{\mathrm{CH}}^{t} = T_{\mathrm{CH}}^{t-\Delta t} + \left(\frac{\left(T_{\mathrm{GDL}}^{t-\Delta t} - T_{\mathrm{CH}}^{t-\Delta t} \right) k_{\mathrm{GDL\_CH}}^{\mathrm{eff}}}{\left(\frac{\delta_{\mathrm{CH}}}{2} + \frac{\delta_{\mathrm{GDL}}}{2} \right) \delta_{\mathrm{CH}}} - \frac{\left(T_{\mathrm{CH}}^{t-\Delta t} - T_{\mathrm{BP}}^{t-\Delta t} \right) k_{\mathrm{BP\_CH}}^{\mathrm{eff}}}{\left(\frac{\delta_{\mathrm{BP}}}{2} + \frac{\delta_{\mathrm{CH}}}{2} \right) \delta_{\mathrm{CH}}} + S_{\mathrm{T}} \right) \frac{\Delta t}{\rho c_{\mathrm{p}}} \tag{8}$$

The heating or cooling effects of inlet gases are considered since inlet temperature may differ from cell temperature.

$$Q_{\mathrm{flow}} = \frac{\sum_{\mathrm{i}} \left(c_p u_{\mathrm{in}} c_{\mathrm{in}} M (T_{\mathrm{CH}} - T_{\mathrm{in}}) \right)_{\mathrm{i}} A_{\mathrm{CH,in}}}{A_{\mathrm{cell}} \delta_{\mathrm{CH}}} (\mathrm{i} = \mathrm{H_2}, \mathrm{O_2}, \mathrm{N_2}, \mathrm{vp}) \tag{9}$$

The output voltage is calculated based on Tafel equations for simplification.

$$V_{\mathrm{out}} = V_{\mathrm{nernst}} + V_{\mathrm{act}} + V_{\mathrm{ohmic}} + K V_{\mathrm{conc}} \tag{10}$$

where K is the correction factor of concentration loss to better fit experimental data in high current density region.

2.2 Membrane Humidifier Model

The schematic diagram of one-dimensional transient planar membrane humidifier is illustrated in Fig. 3. Hot exhausted gases from PEMFC stack flow in wet channel, while inlet reactant gases flow in dry channel. The countercurrent flow is adopted since

it is more convenient for system arrangement. The water contents at two sides of membrane are influenced by water vapor concentration in flow channel, and phase changes occur owing to the difference of membrane water content and equilibrium water content. However, phase changes are seldom taken into consideration in previous numerical studies [18–20] which are actually essential in investigating the dynamic response. The governing equations include membrane water, water vapor, and energy conservation, which are also updated at each time step using the same explicit formulation calculation. The membrane humidifier properties are given in Table 2.

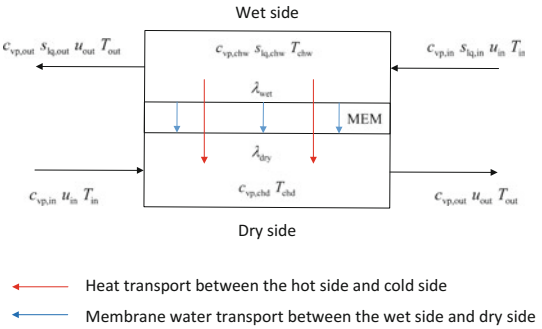


Fig. 3. Schematic diagram of the transient planar membrane humidifier model.

Table 2. Membrane humidifier properties.

Parameter	Value
Membrane area	100 cm^2
Channel length; width; depth; rib width	100; 1.0; 1.0; 1.0 mm
Thickness of membrane	0.025 mm
Heat transfer coefficient between humidifier and surrounding	h_{surr}= 20 W m^{-2} K^{-1}

The membrane water content at two sides of membrane is calculated as

$$\lambda_{wet}^{t} = \lambda_{wet}^{t-\Delta t} + \left(-\frac{\left(\lambda_{wet}^{t-\Delta t} - \lambda_{dry}^{t-\Delta t}\right) D_{m}^{eff}}{\left(\frac{\delta_{MEM}}{2}\right)\frac{\delta_{MEM}}{2}} + S_{MH,mw}\frac{EW}{\rho_{MEM}} \right) \Delta t \tag{11}$$

The water vapor concentrations in channel are calculated as

$$c_{vp,wet}^{t} = c_{vp,wet}^{t-\Delta t} + \left(S_{MH,vp}\right)\Delta t \tag{12}$$

The liquid water volume fraction in wet channel is calculated as

$$s_{\mathrm{lq,chw}}^{t} = s_{\mathrm{lq,chw}}^{t-\Delta t} + \left(S_{\mathrm{MH,lq}}\right)\frac{\Delta t}{\rho_{\mathrm{lq}}} \tag{13}$$

$$S_{\mathrm{MH,lq}} = S_{\mathrm{v-l}} + \frac{\left(s_{\mathrm{lq,in}}u_{\mathrm{in}} - s_{\mathrm{lq,out}}u_{\mathrm{out}}\right)u_{\mathrm{ratio}}\rho_{\mathrm{lq}}A_{\mathrm{MH,in}}}{r_{\mathrm{CH,MH}}A_{\mathrm{MH}}\delta_{\mathrm{CH}}} \tag{14}$$

It should be noted that liquid water in dry channel is not considered since the water vapor concentration is hard to reach saturation.

The temperatures at the center of channels and membrane are calculated as

$$T_{\mathrm{CHw}}^{t} = T_{\mathrm{CHw}}^{t-\Delta t} + \left(-\frac{\left(T_{\mathrm{CHw}}^{t-\Delta t} - T_{\mathrm{MEM}}^{t-\Delta t}\right)k_{\mathrm{CHw\_MEM}}^{\mathrm{eff}}}{\left(\frac{\delta_{\mathrm{CHw}}}{2} + \frac{\delta_{\mathrm{MEM}}}{2}\right)\delta_{\mathrm{CHw}}} - \frac{h_{\mathrm{surr}}\left(T_{\mathrm{CHw}}^{t-\Delta t} - T_{\mathrm{surr}}\right)}{\delta_{\mathrm{CHw}}} + S_{\mathrm{T}}\right)\frac{\Delta t}{\rho c_{\mathrm{p}}} \tag{15}$$

$$T_{\mathrm{MEM}}^{t} = T_{\mathrm{MEM}}^{t-\Delta t} + \left(\frac{\left(T_{\mathrm{CHw}}^{t-\Delta t} - T_{\mathrm{MEM}}^{t-\Delta t}\right)k_{\mathrm{CHw\_MEM}}^{\mathrm{eff}}}{\left(\frac{\delta_{\mathrm{CHw}}}{2} + \frac{\delta_{\mathrm{MEM}}}{2}\right)\delta_{\mathrm{MEM}}} - \frac{\left(T_{\mathrm{MEM}}^{t-\Delta t} - T_{\mathrm{CHd}}^{t-\Delta t}\right)k_{\mathrm{MEM\_CHd}}^{\mathrm{eff}}}{\left(\frac{\delta_{\mathrm{MEM}}}{2} + \frac{\delta_{\mathrm{CHd}}}{2}\right)\delta_{\mathrm{MEM}}} + S_{\mathrm{T}}\right)\frac{\Delta t}{\rho c_{\mathrm{p}}} \tag{16}$$

2.3 Air Compressor Model

The centrifugal air compressor model consists of two parts: characteristics of air compressor and inertance of drive motor. The mass flow rate of air compressor is determined by rotating speed and pressure ratio which is defined as the ratio between outlet pressure and inlet pressure. The compressor rotor is actuated by drive motor and the rotating speed is controlled by motor terminal voltage. The Rotrex C15-16 centrifugal air compressor is adopted in the study. It should be noted that air compressor MAP from supplier is tested at specific conditions which may vary from real operating conditions, making it necessary to modify the MAP for better accuracy.

The relationship among mass flow rate, pressure ratio, and rotating speed is derived by curve fitting tool in MATLAB which can also be achieved by Excel or Origin, etc. To improve polynomial fitting accuracy, it is helpful to transform the predictors by normalizing their center and scale based on sample points. The curve fitting is conducted by binary polynomials of 5 and the fitting results are given as follows:

$$x = \frac{N_{\mathrm{cp}} - 1.444 \times 10^{5}}{4.808 \times 10^{4}}, y = \frac{p_{\mathrm{ratio}} - 1.726}{0.4504} \tag{17}$$

$$\begin{aligned} m_{\mathrm{air}} = {} & \mathrm{p}_{00} + \mathrm{p}_{10}x + \mathrm{p}_{01}y + \mathrm{p}_{20}x^2 + \mathrm{p}_{11}xy + \mathrm{p}_{02}y^2 + \mathrm{p}_{30}x^3 + \mathrm{p}_{21}x^2y + \mathrm{p}_{12}xy^2 + \mathrm{p}_{03}y^3 \\ & + \mathrm{p}_{40}x^4 + \mathrm{p}_{31}x^3y + \mathrm{p}_{22}x^2y^2 + \mathrm{p}_{13}xy^3 + \mathrm{p}_{04}y^4 + \mathrm{p}_{50}x^5 + \mathrm{p}_{41}x^4y + \mathrm{p}_{32}x^3y^2 + \mathrm{p}_{23}x^2y^3 + \mathrm{p}_{14}xy^4 + \mathrm{p}_{05}y^5 \end{aligned} \tag{18}$$

The SSE (the sum of squares due to error) of fitting curve is 0.008865, and R-square (coefficient of determination) is 0.9805, indicating good curve fitting results. However, the surging region which causes damage to air compressor and maximum flow rate region which exceeds compressor ability are included, making it necessary to exclude the aforementioned unfavorable regions. Using the same method, the surging line and maximum flow rate line are obtained.

$$p_{\text{surging}} = 1.009 \times 10^4 m_{\text{air}}^3 + 264.5 m_{\text{air}}^2 + 2.469 m_{\text{air}} + 1.032 \tag{19}$$

$$p_{\text{maxrate}} = 445.9 m_{\text{air}}^3 - 68.58 m_{\text{air}}^2 + 5.975 m_{\text{air}} + 0.9028 \tag{20}$$

The inertance of drive motor is calculated as

$$J_{\text{cp}} \frac{d\omega_{\text{cp}}}{dt} = \tau_{\text{cm}} - \tau_{\text{cp}} \tag{21}$$

Since the rotating speed of drive motor is controlled by motor terminal voltage, the change of voltage determines the mass flow rate and its dynamic response. A simple proportion–integral–derivative (PID) control is adopted to reach the required air flow rate of PEMFC stack.

The drive torque is equal to load torque when rotating speed of compressor stays unchanged, which is expressed as

$$\frac{P_{\text{cp}}}{\omega_{\text{cp}}} = \eta_{\text{cm}} \frac{\kappa_{\text{t}}}{R_{\text{cm}}} (v_{\text{cm}} - \kappa_{\text{v}} \omega_{\text{cm}}) \tag{22}$$

The real air flow rate is detected by flow sensor in supply manifold and sent back to control center. The variance of real mass flow rate and required mass flow rate is set as control deviation, calculating the control deviation and corresponding motor terminal.

2.4 Cooling Model

The coolant flows through bipolar plate and absorbs heat from high-temperature stack, and then it releases heat to surroundings through radiator. Besides, hot air after compressor also needs to be cooled down before it enters cathode membrane humidifier. The amount of heat absorbed by coolant in stack is calculated as

$$Q_{\text{cool}} = c_{\text{p}} \rho V (T_{\text{out}} - T_{\text{in}}) \tag{23}$$

where $V(\text{m}^3\ \text{s}^{-1})$ is the volume flow rate of coolant. The amount of heat taken away by coolant can also be calculated in a simple way using convective heat transfer between stack and coolant.

$$Q_{\text{cool}} = h_{\text{cool}} A_{\text{cool}} (T_{\text{BP}} - T_{\text{cool}}) \tag{24}$$

The amount of heat which needs to be absorbed by coolant after air compression is calculated as

$$Q_{\mathrm{cp,air}} = c_{\mathrm{p}} m_{\mathrm{air,cp}} \left(T_{\mathrm{cp,out}} - T_{\mathrm{set}}\right) \tag{25}$$

The amount of heat released to surroundings by cooling fan is calculated as

$$Q_{\mathrm{fan}} = c_{\mathrm{p}} \rho_{\mathrm{air}} V_{\mathrm{air}} \left(T_{\mathrm{air,out}} - T_{\mathrm{air,in}}\right) \tag{26}$$

Since the heat carried by coolant is released by radiator, the following equation can be obtained:

$$Q_{\mathrm{rad}} = Q_{\mathrm{cool}} + Q_{\mathrm{cp,air}} \tag{27}$$

Strictly speaking, the heat convection among coolant, PEMFC stack, and cooling air should be calculated more precisely which needs the overall heat transfer coefficient. The heat transfer coefficient, temperature differences of coolant after flowing through PEMFC stack, and radiator are artificially given in a reasonable range.

2.5 Boundary and Initial Conditions

In the study, constant current density mode is adopted, indicating that the total current of single cell is the same among PEMFC stack. However, local current density from upstream to downstream in flow direction may vary since reactant concentrations keep decreasing when it flows toward channel outlet. Since bipolar plate has good electrical conductivity, it is more reasonable to assume that the voltage among different nodes (shown in Fig. 2) in single cell is the same rather than assuming uniform local current density along gas channel. There is a consistent one-to-one match between output voltage and current density, making it feasible to obtain local current density by solving the system of nonlinear equations.

$$\begin{cases} V_{\mathrm{nernst}} + V_{\mathrm{act}}(I_1) + V_{\mathrm{ohmic}}(I_1) + V_{\mathrm{conc}}(I_1) = V_{\mathrm{out}} \\ V_{\mathrm{nernst}} + V_{\mathrm{act}}(I_2) + V_{\mathrm{ohmic}}(I_2) + V_{\mathrm{conc}}(I_2) = V_{\mathrm{out}} \\ \qquad\qquad\qquad\qquad \vdots \\ V_{\mathrm{nernst}} + V_{\mathrm{act}}(I_n) + V_{\mathrm{ohmic}}(I_n) + V_{\mathrm{conc}}(I_n) = V_{\mathrm{out}} \\ I_1 + I_2 + \qquad\qquad \cdots \qquad\qquad I_{n-1} + I_n = n \cdot I_{\mathrm{set}} \end{cases} \tag{28}$$

The Newton–Raphson method is adopted to solve the $n + 1$ variables in Eqs. (5–1), obtaining local current density of all nodes in flow direction.

Only convective heat loss from end bipolar plates to environment in the through-plane direction is considered in PEMFC stack.

$$Q = h_{\mathrm{surr}} A (T_{\mathrm{BP}} - T_{\mathrm{surr}}) \tag{29}$$

In order to guarantee sufficient reactant gas humidification, the membrane humidifier temperature may be much higher than surroundings, making it necessary to consider the heat loss.

$$Q_{MH} = h_{surr} A_{MH} (T_{MH} - T_{surr}) \tag{30}$$

For simplification, pressure drop in air filter and supply manifold are neglected. The inlet temperature of air compressor is the same as surroundings and the inlet pressure is atmospheric pressure. The inlet temperature of cathode membrane humidifier dry channel is the temperature set by radiator. The inlet gases in wet channel are exhausted gases from PEMFC stack. The inlet hydrogen flow rate of stack is supplied based on anode stoichiometry, consisting of pure hydrogen from high-pressure tank and unused hydrogen mixed in anode exhausted gases.

The initial cell temperature is set as surrounding temperature, and liquid water/ice volume fraction in porous layers is zero. The explicit calculation equations are updated with a time step of 10^{-6} s to guarantee accuracy and calculation efficiency simultaneously.

3 Results and Discussion

3.1 Model Validation

The various models have been carefully compared with experimental data in the literature. As regards proton exchange membrane fuel cell (PEMFC) stack model, the grid independence study is conducted at first to show the effects of node number in flow channel on mass transport processes and output performance. It is found that relative error decreases with growing node number and five nodes are enough to guarantee calculation accuracy. Further increasing node number contributes negligibly but slows down calculation efficiency. Then, the PEMFC stack model has been compared with experimental data of single cell [24] and fuel cell stack [25] under various operating conditions as shown in Fig. 4a, b. The planar membrane humidifier model has also been compared with experimental data [26] which is shown in Fig. 4c, d. Besides, the stable and dynamic performance of air compressor model under aforementioned PID control strategy has been investigated, which shows good accuracy and dynamic response.

3.2 Cold Start Performance

The cold start performances from two typical temperatures, −10 °C and −5 °C, are investigated. Larger initial water contents represent less purge duration while lower initial contents result from longer purge duration. The same initial water contents in fuel cell and membrane humidifier refer to the cases that gas purge happens in both of them. Keeping the initial water content in humidifier constant while changing that in fuel cell represents the cases that gas purge merely occurs in fuel cell stack, implying there exists an additional purge pipeline and the purging gas does not flow through

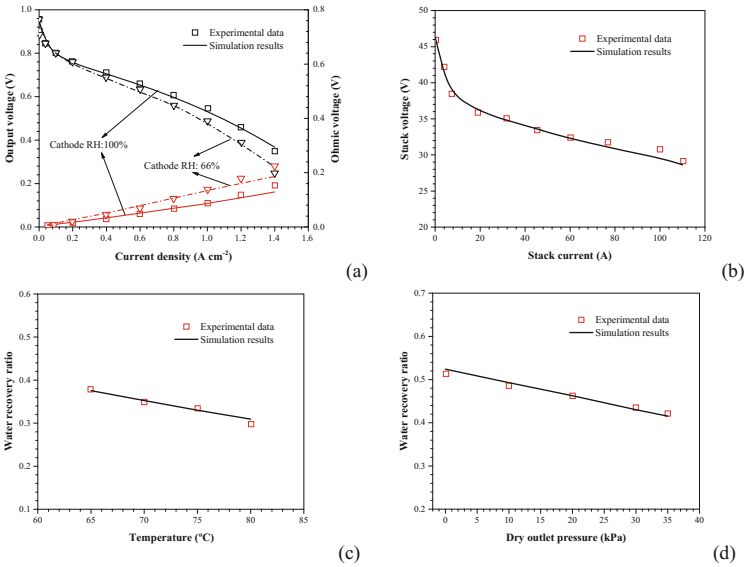


Fig. 4. Comparison between simulation results and experimental data. (a) Single cell [24]. (b) Stack [25]. (c) Membrane humidifier [26]. (d) Membrane humidifier [26].

humidifier. Figure 5(a) shows the single cell voltage starting from −10 °C, 0.1 A cm^{-2}. It is found that lower initial membrane content results in longer cold start duration since membrane electrode assembly (MEA) is able to absorb more generated water, delaying ice formation. In the beginning, lower initial water content induces larger output voltage drop owing to larger ohmic voltage loss. There are no apparent differences among the solid lines and dash lines, indicating that membrane humidifier plays little role in cold start performance although the differences are growing with decrease of initial water content in fuel cell. The corresponding stack cathode inlet relative humidity (RH) is demonstrated in Fig. 5(b). The cathode inlet vapor concentration is dependent on humidifier dry channel outlet, while saturated vapor pressure is determined by stack channel temperature. It is observed that higher initial water content in humidifier results in higher stack inlet relative humidity. The inlet RH keeps decreasing since cell temperature is increasing owing to released heat. It should be noticed that curve slopes are almost the same because the amount of heat is mainly decided by current density. Although the inlet RH varies in a large range, the cell voltage shows no apparent differences because saturated vapor pressure in subzero temperature is relatively small. The single cell voltage starting from −5 °C is demonstrated in Fig. 5(c). It is observed that single cell voltage reaches similar value after around 100 s although there are significant voltage differences in the first 50 s. The electroosmotic drag (EOD) effect carries water molecules from anode to cathode which decreases hydration in anode catalyst layer (CL), resulting in large ion transport resistance in the first 10 s. The ohmic loss declines as more product water are generated and MEA are better hydrated. However, membrane humidifier still has little effect and the output voltage is mainly determined by initial water content in fuel cell stack.

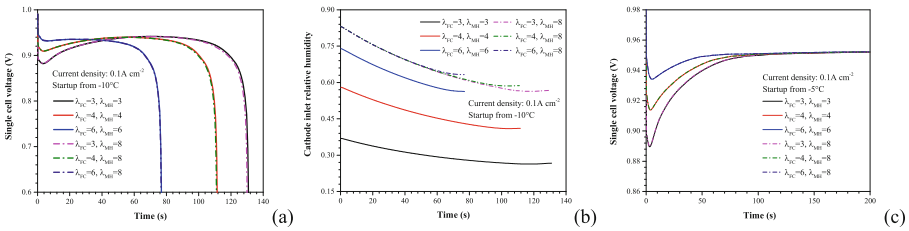


Fig. 5. Output voltage and cathode inlet relative humidity under different purging strategies and cold start temperatures. (a) Single cell voltage, startup from −10 °C. (b) Cathode inlet relative humidity, startup from −10 °C. (c) Single cell voltage, startup from −5 °C.

Figure 6a shows the local current density distribution in flow direction of a single cell starting from −10 °C, 0.1 A cm^{-2}, $\lambda_{FC} = \lambda_{MH} = 6$. The node 1 has the highest current density while node 5 has the lowest current density in the first 60 s. The reactant gas concentrations at fuel cell inlet are higher than outlet since hydrogen and oxygen are consumed all the way and higher reactant gas concentrations lead to less voltage loss. It should be noticed that ice volume fraction in node 1 is always higher than that in node 5 because more supercooled water is generated, making it harder for upstream reactant gases to reach CL. In the last 15 s, node 5 possesses the highest local current density, while node 1 possesses the lowest current density. The activation loss and concentration loss are smaller in downstream than that in upstream since less ice is formed. The voltage difference is thus compensated by ohmic voltage loss and this leads to higher local current density in downstream since output voltage of all nodes is assumed the same.

The membrane water content in cathode CL is shown in Fig. 6b. The membrane water content experiences a quick increase in the first 6 s, and then it rises slowly because membrane water turns into supercooled water and vapor simultaneously after it gets saturated. It is observed that node 1 also has the highest water content, while node 5 has the lowest water content in the first 70 s. Although water vapor concentration in downstream is usually higher than that in upstream, more membrane water is generated in upstream due to higher local current density. Figure 6c shows the temperature distribution in cathode CL. It is observed that temperature in upstream is higher although inlet reactant gas temperature is always lower than fuel cell stack which

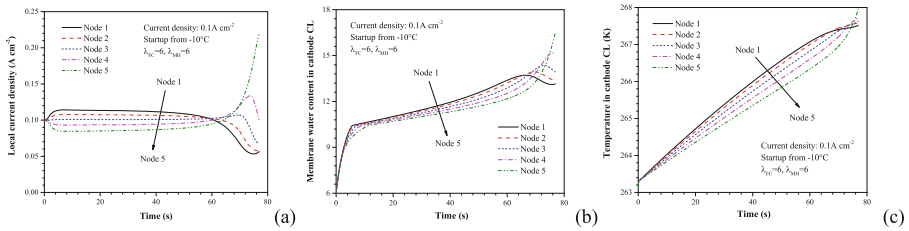


Fig. 6. Current density, membrane water content, and temperature distribution in flow direction starting from −10 °C, 0.1 A cm^{-2}, $\lambda_{FC} = \lambda_{MH} = 6$. (a) Local current density. (b) Membrane water content in cathode CL. (c) Temperature in cathode CL.

actually acts as cooling sources. Since membrane humidifier is heated by stack exhausted gases, the temperature of inlet reactant gases is slightly higher than surrounding temperature.

3.3 Normal Startup Performance

Effects of gas purge duration are also investigated when stack is started from normal temperature like 30 °C. Figure 7(a) shows the single cell voltage under different initial membrane water contents in fuel cell and humidifier. The voltage experiences a sharp drop in the first 2 s, especially in the case with relatively low initial membrane water content. Then, the voltage increases owing to better MEA hydration and reaches a relatively stable value after 25 s since cell temperature is kept at around 60 °C with the help of cooling subsystem. The cases with higher initial water content in membrane humidifier demonstrate higher output voltage than the cases with lower value during startup period. The slight voltage fluctuation before stabilization is attributed to excessive liquid water presence and membrane water content oversaturation. In addition to large amount of water generated in high current density like 1.2 A cm^{-2}, membrane humidifier also supplies a lot of water vapor. The large residual liquid water in CL hinders the phase changes among membrane water, vapor, and liquid water, resulting in temporary membrane water oversaturation. After liquid water is driven out of porous layers, the membrane water content in MEA returns to equilibrium and the single cell voltage also declines (shown in the enlarged graph of Fig. 7a). To avoid sharp voltage drop during startup process, long purge duration is not suggested for both fuel cell stack and membrane humidifier after shutdown. Although the initial water content in membrane humidifier varies, it reaches the same value at last as shown in Fig. 7b. The membrane water content in wet side is always higher than that in dry side since exhausted hot wet gases from PEMFC stack flow into wet channel, while dry gases flow into dry channel. Besides, it takes shorter time to reach stable value with higher initial water content in humidifier while initial water content in fuel cell stack has little effect. The water content of dry side experiences a slight drop in the beginning because the amount of air which flows into dry channel is larger than that flows into wet channel. Thus, the humidifier membrane water has to compensate for vapor amount difference. However, the exhausted gases grow hotter and wetter as fuel cell temperature is increased from 30 °C to 60 °C which enlarges the amount of wet channel inlet vapor, making it possible for humidifier to absorb and store water from exhausted gases. The relative humidity at stack cathode inlet and humidifier dry side outlet is shown in Fig. 7c. It is found that RH curves follow the same trend as membrane water content. The noteworthy variations between stack inlet RH and humidifier outlet RH are caused by temperature difference between stack and humidifier. The highest local temperature in PEMFC stack is kept at around 60 °C while the humidifier temperature is stabilized at around 42 °C as a result of exhausted gases heating and convective heat loss to environment. The stack inlet RH is almost the same, which also accounts for the little difference of single cell voltage.

The membrane humidifier performance under various stack current densities is demonstrated in Fig. 8. With rise of current density, membrane water content in humidifier increases more rapidly since more water vapor is generated in PEMFC stack

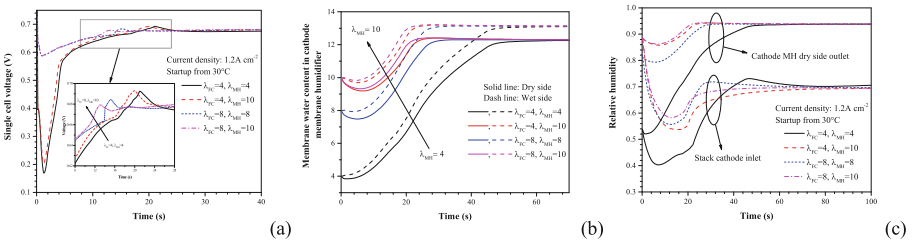


Fig. 7. Single cell voltage, water content, and relative humidity under different purging strategies, startup from 30 °C, 1.2 A cm^{-2}. (a) Single cell voltage. (b) Membrane water content. (c) Relative humidity °C.

and flows into humidifier. Meanwhile, the membrane humidifier is better heated with more hot exhausted gases, leading to higher temperature as shown in Fig. 8(b). However, higher temperature means higher saturated vapor pressure, and it results in relatively smaller stabilized water content value. It should be noticed that membrane water content drop in the first 20 s decreases with larger current density owing to the rising amount of water vapor which flows into humidifier.

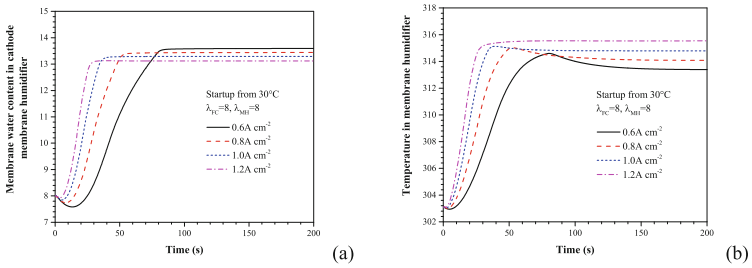


Fig. 8. Membrane humidifier performance under different stack current densities, startup from 30 °C, $\lambda_{FC} = \lambda_{MH} = 8$. (a) Membrane water content in wet side channel. (b) Temperature.

4 Conclusion

In the study, a comprehensive proton exchange membrane fuel cell (PEMFC) system is developed, including a two-dimensional transient multiphase stack model, a transient membrane humidifier model, air compressor model, and cooling model. The coupled water and heat transport processes, water phase changes, gas/liquid transport in porous layers, and flow channel are taken into consideration. All models have been individually validated against experimental data which show good agreement. Effects of gas purge duration in PEMFC stack and membrane humidifier on startup performance are investigated under subzero and normal temperatures. It is found that purge duration of humidifier has little effect on stack output voltage when started from −10 °C and −5 °C since saturated vapor pressure is small in subzero temperatures. The cold start duration is mainly determined by initial membrane water content in PEMFC stack and a slight voltage drop is observed in the beginning owing to increased ionic resistance in

membrane electrode assembly (MEA). The local current density in upstream is usually higher because reactant gases are more abundant. Higher current density results in more generated water and heat, leading to higher membrane water content and temperature. To avoid sharp voltage drop during startup from 30 °C, long purge duration is not suggested for both fuel cell stack and membrane humidifier. The noteworthy variations between stack inlet relative humidity (RH) and humidifier outlet RH are caused by temperature difference between stack and humidifier. The stack temperature can be maintained at around 60 °C with cooling subsystem, while the humidifier temperature is stabilized at about 42 °C as a result of exhausted gases heating and heat loss to environment. The membrane water content in humidifier increases more rapidly with rise of current density since more water vapor is generated and flows into humidifier, resulting in higher membrane humidifier temperature.

Acknowledgements. The research is supported by the National Key Research and Development Program of China (2018YFB0105505) and National Natural Science Foundation of China for Excellent Young Scholars (51622606).

References

1. Rabbani, A., Rokni, M.: Dynamic characteristics of an automotive fuel cell system for transitory load changes. Sustain. Energy Technol. Assess. **1**, 34–43 (2013)
2. Guo, Y.F., Chen, H.C., Wang, F.C.: The development of a hybrid PEMFC power system. Int. J. Hydrog. Energy **40**(13), 4630–4640 (2015)
3. Hosseinzadeh, E., Rokni, M., et al.: Thermal and water management of low temperature proton exchange membrane fuel cell in fork-lift truck power system. Appl. Energy **104**, 434–444 (2013)
4. Santamaria, A.D., Bachman, J., Park, J.W.: Cold-start of parallel and interdigitated flow-field polymer electrolyte membrane fuel cell. Electrochim. Acta **107**, 327–338 (2013)
5. Yang, Z., Du, Q., Huo, S., et al.: Effect of membrane electrode assembly design on the cold start process of proton exchange membrane fuel cells. Int. J. Hydrog. Energy **42**(40), 25372–25387 (2017)
6. Xie, X., Zhang, G., Zhou, J., et al.: Experimental and theoretical analysis of ionomer/carbon ratio effect on PEM fuel cell cold start operation. Int. J. Hydrog. Energy **42**(17), 12521–12530 (2017)
7. Fan, L., Zhang, G., Jiao, K.: Characteristics of PEMFC operating at high current density with low external humidification. Energy Convers. Manag. **150**, 763–774 (2017)
8. Jiao, K., Li, X.: Water transport in polymer electrolyte membrane fuel cells. Prog. Energy Combust. Sci. **37**(3), 221–291 (2011)
9. Tabe, Y., Yamada, K., Ichikawa, R., et al.: Ice formation processes in PEM fuel cell catalyst layers during cold startup analyzed by cryo-SEM. J. Electrochem. Soc. **163**(10), F1139–F1145 (2016)
10. Niya, S.M.R., Hoorfar, M.: Study of proton exchange membrane fuel cells using electrochemical impedance spectroscopy technique—a review. J. Power Sources **240**, 281–293 (2013)
11. Ishikawa, Y., Shiozawa, M., Kondo, M., et al.: Theoretical analysis of supercooled states of water generated below the freezing point in a PEFC. Int. J. Heat Mass Transf. **74**, 215–227 (2014)

12. Plazanet, M., Sacchetti, F., Petrillo, C., et al.: Water in a polymeric electrolyte membrane: sorption/desorption and freezing phenomena. J. Membr. Sci. **453**, 419–424 (2014)
13. Molaeimanesh, G.R., Akbari, M.H.: Impact of PTFE distribution on the removal of liquid water from a PEMFC electrode by Lattice Boltzmann method. Int. J. Hydrog. Energy **39**(16), 8401–8409 (2014)
14. Niu, Z., Jiao, K., Zhang, F., et al.: Direct numerical simulation of two-phase turbulent flow in fuel cell flow channel. Int. J. Hydrog. Energy **41**(4), 3147–3152 (2016)
15. Srinophakun, T., Martkumchan, S.: Ionic conductivity in a Chitosan membrane for a PEM fuel cell using molecular dynamics simulation. Carbohydr. Polym. **88**(1), 194–200 (2012)
16. Cheng, S.J., et al.: Investigating the effects of operational factors on PEMFC performance based on CFD simulations using a three-level full-factorial design. Renew. Energy **39**(1), 250–260 (2012)
17. Luo, Y., Guo, Q., Du, Q., et al.: Analysis of cold start processes in proton exchange membrane fuel cell stacks. J. Power Sources **224**, 99–114 (2013)
18. Yu, S., et al.: A parametric study of the performance of a planar membrane humidifier with a heat and mass exchanger model for design optimization. Int. J. Heat Mass Transf. **54**(7–8), 1344–1351 (2011)
19. Bhatia, D., Sabharwal, M., Duelk, C.: Analytical model of a membrane humidifier for polymer electrolyte membrane fuel cell systems. Int. J. Heat Mass Transf. **58**(1–2), 702–717 (2013)
20. Park, S., Jung, D.: Effect of operating parameters on dynamic response of water-to-gas membrane humidifier for proton exchange membrane fuel cell vehicle. Int. J. Hydrog. Energy **38**(17), 7114–7125 (2013)
21. Tajiri, K., Tabuchi, Y., et al.: Effects of operating and design parameters on PEFC cold start. J. Power Sources **165**(1), 279–286 (2007)
22. Tang, H.Y., Santamaria, A.D., Bachman, J., et al.: Vacuum-assisted drying of polymer electrolyte membrane fuel cell. Appl. Energy **107**, 264–270 (2013)
23. Du, Q., Jia, B., Luo, Y., et al.: Maximum power cold start mode of proton exchange membrane fuel cell. Int. J. Hydrog. Energy **39**(16), 8390–8400 (2014)
24. Ozen, D.N., et al.: Effects of operation temperature and reactant gas humidity levels on performance of PEM fuel cells. Renew. Sustain. Energy Rev. **59**, 1298–1306 (2016)
25. Hwang, J.J.: Effect of hydrogen delivery schemes on fuel cell efficiency. J. Power Sources **239**, 54–63 (2013)
26. Kadylak, D., Mérida, W.: Experimental verification of a membrane humidifier model based on the effectiveness method. J. Power Sources **195**(10), 3166–3175 (2010)

Preliminary Simulation Study of Flow Field Around a Spark Plug Under Ambient and Engine Conditions

Navjot Sandhu, Shouvik Dev, Divyanshu Purohit, Zhenyi Yang, Ming Zheng(✉), and David Ting

Department of Mechanical, Automotive and Materials Engineering, University of Windsor, 401 Sunset Ave, Windsor, ON, Canada
mzheng@uwindsor.ca

Abstract. Sustainable transportation solutions for the future would require advanced powertrains which can meet the goals of emission and fuel consumption reduction. One option could be high-efficiency spark ignition (SI) internal combustion engines using conventional or renewable fuels. Such SI engines in the future may operate under lean conditions at which the air is in excess with respect to the fuel. Typically, ignition and complete combustion of such a lean mixture of air and fuel is a challenge owing to the reduced charge reactivity. One solution is to enhance the in-cylinder charge motion to increase the flame velocity. However, this charge motion can affect the initial spark breakdown and the consequent flame kernel development. Therefore, in order to estimate the flow field around the spark plug, a simulation study is undertaken. The simulations are performed using Converge™ three-dimensional simulation suite (version 2.3). ANSYS EnSight (version 10.1) is used for post-processing of the simulation data. Two types of flow fields are simulated. The first flow field simulates a cross-flow of air across the electrode gap of a conventional J-type spark plug under ambient pressure. The flow upstream of the plug is laminar, and the flow velocity is varied. This part of the study is used to determine the effect of the spark plug geometry on the flow. The second condition simulates the in-cylinder flow field of a two-valve, single-cylinder engine. The intake air flow rate is the main variable. The numerical estimation of the flow velocity and the turbulence around the spark gap are correlated with the experimental results. Preliminary results indicate that the spark plug can generate turbulence in its wake under steady flow conditions and the vorticity magnitudes can be correlated to the electrical parameters. In the engine, the flow field in and near the spark gap may not correlate with the bulk air motion.

1 Introduction

The quest for sustainable transportation solutions has prompted research and development into advanced powertrains that can meet or exceed standards for emissions and fuel efficiency. Such advanced powertrains may continue to use spark ignition (SI) internal combustion engines with conventional or low-carbon footprint alternative fuels [1–3]. Since gasoline-fueled SI engines dominate the international passenger

A. Vasel and D. S.-K. Ting (Eds.): EAS 2018, SPE, pp. 34–54, 2019.
https://doi.org/10.1007/978-3-030-00105-6_3

vehicle market, improvements in efficiency can make a significant impact on energy security and environment [4]. Lean and diluted combustion have been proposed as two effective methods for increasing the efficiency of SI engines and decreasing the engine-out emissions [5, 6]. Conventional SI engines typically operate under stoichiometric conditions with all the intake air being utilized for complete combustion of the fuel. For lean combustion, intake air supply is more than what is required for complete combustion of the fuel. For diluted combustion, the fresh intake air is mixed with the recirculated exhaust gas (process referred to as exhaust gas recirculation (EGR)) before this mixture is introduced into the cylinder. In addition to reducing nitrogen oxides (NOx) emissions due to lower flame temperature, EGR also has a positive impact on fuel consumption due to reduced throttling losses [7–10]. However, operating the SI engine under lean conditions or with EGR dilution reduces the reactivity of the fuel–air mixture which could lead to poor ignitability. This in turn would limit the level of dilution due to the onset of misfire [5]. Another adverse consequence of lean or diluted combustion is the lower flame speed [11]. The lower flame speed could reduce the fuel efficiency due to incomplete combustion of the fuel–air mixture.

One of the strategies to counter these effects and accelerate the flame velocity is to increase turbulence inside the cylinder by enhancing the motion of the fuel–air mixture. It has been established through experimental and numerical studies that the fluid flow characteristics of the in-cylinder mixture affect the combustion process at a fundamental level [12]. Faster flame speed in an SI engine can help in boosting the thermal efficiency and reducing the NOx emissions by extending the stable lean burn limit, thus making very diluted fuel–air mixtures viable for engine operation [13]. However, this increased in-cylinder turbulence can have adverse effects on the initial spark breakdown and the subsequent flame kernel development process. Excess turbulence can also lead to extinguishment of the flame. Instability in the early stages of combustion can significantly contribute toward cyclic variations in the engine operation [14–16]. Studies have shown that the plasma channel in the spark gap tends to stretch and follow the flow field [17, 18]. While this stretch helps in increasing the total surface area of the spark plasma, the energy density of the spark channel can diminish to a large extent, depending on the flow velocity. Consequently, the air–fuel mixture can either fail to ignite entirely or the initial flame kernel can form and then later be extinguished due to the inability to sustain the combustion [17]. In order to mitigate the negative effects of turbulence on the formation and sustainability of the spark channel, it is imperative to understand the flow field around the spark plug during ignition timing. Eventually, this flow field may be controlled to optimize the combustion event.

There has been a substantial amount of research on the effects of air motion on the spark ignition using constant volume combustion chambers as well as research engine platforms [16–22]. Sayama et al. studied the spark ignition and early flame development at lean air–fuel ratios of 20–30 under high-velocity flow conditions of $\sim$65 m/s at the spark gap in a swirl chamber. The authors found that the spark channels stretched downstream of the spark plug to follow the flow field and frequent restrikes or short circuits were apparent [17]. The findings also demonstrated that the degree of dilution had a significant effect on flame development and sustainability following the initial spark event. Aleiferis et al. used a single-cylinder optical research engine with a pent-roof combustion chamber and variable valve actuation to study the effect of flow

characteristics and spark plug orientation on early flame kernel development and cyclic variations [18]. Direct imaging demonstrated the effect of orientation of spark plug with respect to the flow on the spark channel stretch and the subsequent cyclic variations in the flame development. Additionally, they observed that the flame kernel shape was never spherical and the aspect ratio indicated the dominant effect of large-scale turbulence features. Johansson investigated the effect of several important parameters including flow field, gas composition, and temperature on the early combustion [16]. He concluded that substantial amount of fluctuations during early combustion at high engine load conditions can be attributed to the flow field turbulence and the mixture composition near the spark plug.

Arcoumanis and Bae studied ignition of propane–air mixtures using laser Doppler velocimetry (LDV) and shadowgraphy in a constant volume combustion vessel [19]. They reported that the orientation of spark plug with respect to the mean flow had greater effect at higher velocity as compared to low-velocity cases. Plug orientation with ground electrode downstream of the flow had the highest flame velocity, whereas the upstream ground electrode configuration dampened the effect of turbulence and slowed down the flame speed possibly due to shrouding of the flow by the ground electrode. Ballal and Lefebvre investigated the effect of air flow on electrical parameters of spark ignition in a specially designed wind tunnel at different pressures and velocities of up to 100 m/s [20]. Their findings demonstrated that spark duration decreased with increase in velocity, whereas the minimum energy required for successful ignition showed an increasing trend. Yu et al. studied spark discharge affected by directed flow on the spark gap in a constant volume combustion chamber through optical diagnosis and measurement of electrical parameters [21]. The authors observed the stretch of the spark plasma channel underflow with continuous breakage and re-establishment of the channel (termed as restrike process). The frequency of this restrike process increased with increase inflow velocity, thereby highlighting the difficulty in maintaining the spark plasma channel under high flow conditions.

Based on the authors' search, the flow conditions around the spark plug at the time of ignition have a profound effect on the spark discharge and the subsequent formation of the spark flame kernel. Given the difficulty in the ignition of lean or diluted fuel–air mixtures, the importance of the flow structure is of greater significance. Therefore, in the present research, the authors have undertaken a simulation study to estimate the flow field in the vicinity of the spark plug during the spark event. This study is divided into two parts. The first part of the study simulates a simplified condition in which a steady and uniform cross-flow of air is directed across the electrode gap of a conventional J-type spark plug under ambient conditions. The objective is to determine the effect of the spark plug geometry on the flow field around the spark plug. The second part of the study simulates the in-cylinder flow field of a two-valve, single-cylinder engine during the spark timing window. The simulations are performed using Converge^{TM} three-dimensional simulation suite (version 2.3). ANSYS EnSight (version 10.1) is used for post-processing of the simulation data. The flow field is quantified around the spark plug, especially the spark gap, under different throttle openings and engine speeds. The broader objective of this study is to provide the engine operating conditions which could be favorable to initiate the spark discharge.

This paper is organized as follows. The methodology including the engine research platform which is used to validate the in-cylinder flow field model and the details of the CFD simulation parameters are described in the following section. The simulation validation results are presented in the subsequent section followed by the results for the steady cross-flow and the in-cylinder flow field. The final section describes the summary and main conclusions of this research.

2 Research Methodology

2.1 Engine Research Platform

The test engine used in this study was a Yanmar NFD-170, a two-valve single-cylinder stationary diesel engine, which was heavily modified and instrumented for spark ignition research. This included reducing the compression ratio of the diesel engine to 9.2:1 by modifying the piston bowl. The specifications of the engine are summarized in Table 1. The engine was connected to a General Electric® 26G215 direct current (DC) dynamometer. Intake pressure was controlled through an electronic pressure regulator. The intake air flow rate was measured using a ROOTS volumetric flow meter. The spark plug used for ignition was a conventional J-type resistive spark plug with 14 mm metric thread. In-cylinder pressure was recorded using Kistler piezoelectric pressure transducer and was synchronized with the crank position using a crankshaft rotary encoder and camshaft sensor. National Instruments-Data Acquisition (NI-DAQ) card along with the LabVIEW 2010 software package was used to record the pressure data at 0.1-crank angle degree (°CA) resolution. This engine research platform is instrumented for detailed emission measurement. For this study, the engine research platform was only used to determine the motoring pressure traces (no combustion) for validation of the simulated gas exchange process.

Table 1. Engine specifications

Engine	Yanmar NFD-170
Engine type	4-stroke spark ignited
Bore (mm)	102
Stroke (mm)	105
Displacement (cm^3)	858
Connecting rod length (mm)	165
Compression ratio	9.2:1
Intake valve closing (BTDC[a])	135°
Intake valve open interval (CA[b])	229°

[a]Before Top Dead Center
[b]Crank Angle

2.2 CFD Simulation

Two types of flow field were simulated using Converge™ three-dimensional simulation suite. For the first case, a steady and uniform cross-flow of air across the spark gap was

simulated at different flow velocities to investigate the effect of spark plug geometry on the flow field (Fig. 1). A steady-state density-based solver was used. The spark plug was placed in a large volume to ensure that the flow is fully developed before it reaches the spark gap. For the cross-flow case, a constant inlet velocity boundary condition was used. Wall temperature and initial pressure were 298°K and 1 bar, respectively. The base grid size was 4 mm with the fourth level of grid refinement near the spark plug, reducing the grid size in the vicinity of the spark plug to 0.25 mm. The Renormalized Group (RNG) k-ε turbulence model was used to model the turbulence.

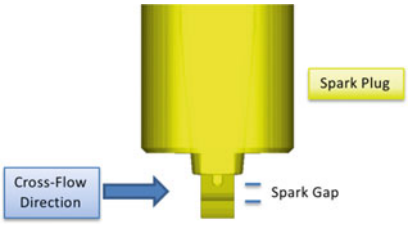


Fig. 1. Spark plug model for cross-flow case

Second case was the engine motoring simulation at different intake mass flow rates and engine speeds. A transient density-based solver was utilized along with the RNG k-ε turbulence model. Since the intake port geometry has a significant influence on the in-cylinder flow [13], the intake port geometry was modeled with high accuracy. To this end, a flexible foam mold cast of the intake port was made and laser scanned. The geometry input for the engine motoring cases is shown in Fig. 2.

The simulation time period extended from 360° BTDC to 100° after TDC. Selective grid refinement was performed to get a balance between computational accuracy and simulation time. Base grid size of 4 mm was used for the model. Third-level grid refinement was performed at the intake valve region during the intake valve event and in the vicinity of the spark plug in the spark discharge window of 100° BTDC to −30° BTDC. Additionally, adaptive mesh refinement (AMR) using velocity was activated in the cylinder region. For the engine motoring case, the intake manifold pressure measurement at 0.1° crank angle resolution was used as a boundary condition for the intake port inflow boundary. The intake and exhaust valve lift profiles were determined experimentally from the test engine and were used as input for the intake and exhaust valve moving boundaries. The wall temperature was 353 K for the boundaries in contact with the in-cylinder gas in accordance with the coolant temperature of the test engine setup. The boundary conditions are listed in Table 2. To account for the crevice volume effect, the effective compression ratio was adjusted to 8.7 for all flow conditions. ANSYS EnSight 10.1 was used to post-process the simulation results data and visualize the flow. The next section describes the validation results.

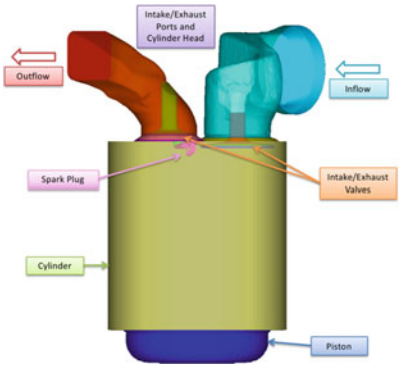


Fig. 2. Geometric model for Engine motoring

Table 2. Engine model boundary conditions

Boundary	Boundary condition
Intake/exhaust port	No-slip stationary wall
Cylinder head	No-slip stationary wall
Inflow	Manifold pressure
Outflow	Zero-gradient pressure
Intake/exhaust valves	No-slip moving wall
Spark plug	No-slip stationary wall
Cylinder	No-slip stationary wall
Piston	No-slip moving wall

3 Validation of Engine Motoring Simulation Gas Exchange Process

The first step in simulating the in-cylinder flow field is to accurately model the gas exchange process. An accurate modeling of the gas exchange process should provide similar profile between the experimental motoring pressure trace and the simulation pressure trace. This comparison is presented in Fig. 3. The experimental and simulated motoring pressures are well correlated for all mass air flow rates.

The mass air flow rates (MAF) of up to 4.8 g/s are achieved under normally aspirated conditions with gradual opening of the throttle. Thereafter, the throttle position is kept constant, and the boost pressure is increased gradually to increase the MAF to 15.1 g/s. Table 3 summarizes the validation conditions and lists the error

Table 3. Validation conditions and errors

MAF [g/s]	p_{int} [kPa gauge]	Δp_{max_err} [%]	MAF [g/s]	p_{int} [kPa gauge]	Δp_{max_err}
3.05	0	−0.02	**7.40**	40	1.30%
4.20	0	−0.29	**10.00**	80	1.78%
4.80	0	−0.24	**15.10**	95	0.49%

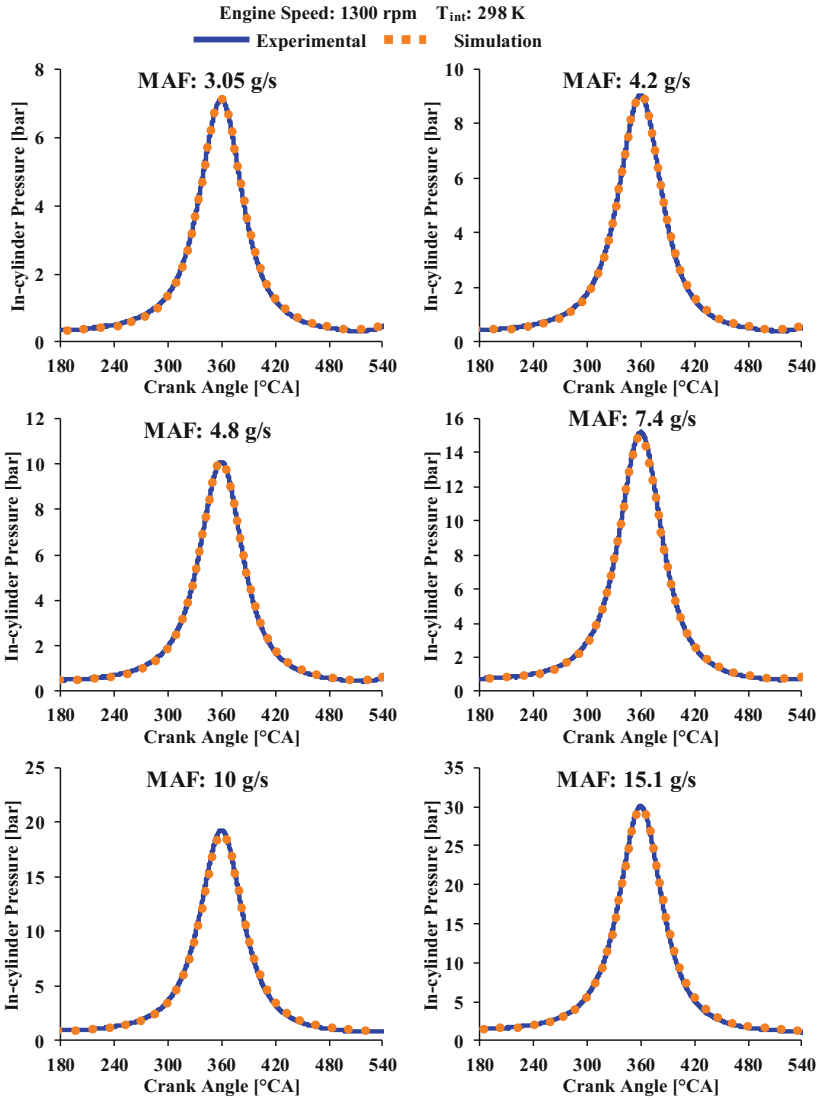


Fig. 3. In-cylinder pressure validation at different mass air flow rates

between the experimental and simulated maximum pressure (denoted by Δp_{max_err}), calculated using Eq. (1) (expressed as percentage). All errors were below 2%.

$$\Delta p_{\max\_err} = \left(p_{\max\_exp} - p_{\max\_sim}\right)/p_{\max\_exp} \tag{1}$$

where p_{max_exp} and p_{max_sim} refer to the experimental and simulated maximum in-cylinder pressures, respectively.

4 Results and Discussion

This section is divided into two subsections. The first subsection describes the results of the cross-flow simulation. The second subsection describes the results of the engine motoring simulation.

4.1 Steady and Uniform Cross-Flow Simulation

The steady cross-flow simulation is performed at inflow velocities of 10, 20, 40, and 100 m/s. The flow conditions in the spark gap are expected to affect the initial breakdown and spark kernel formation. Therefore, the average values of turbulent velocity and turbulence intensity are calculated in the spark gap over a cylindrical volume which is 0.9 mm in diameter and 0.77 mm in height bringing the total volume to approximately 0.5 mm^3.

Figure 4 shows the turbulent velocity and the turbulence intensity in the gap. Turbulent velocity "u'" is defined as the root mean square value of the fluctuating velocity components. Mathematically, it is calculated from the turbulent kinetic energy using expression given by Eq. (2) [23].

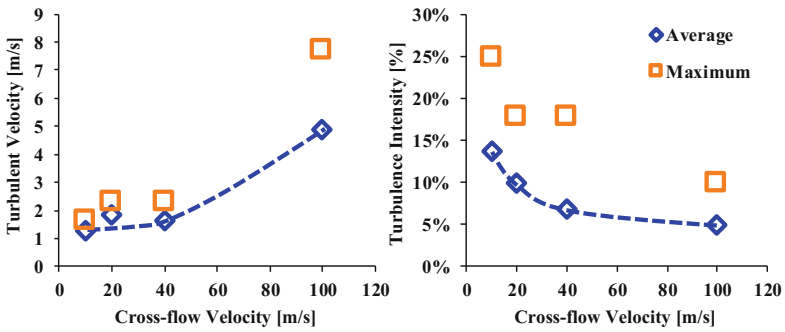


Fig. 4. Steady cross-flow – fluctuating velocity (left) and turbulent intensity (right)

$$u' = \sqrt{\frac{2}{3}k} \tag{2}$$

where "k" represents the turbulent kinetic energy of the flow. Turbulence intensity "I" is the ratio of turbulent velocity to the mean flow velocity magnitude "$\overline{U}$" and is given by Eq. (3).

$$I = \frac{u'}{\overline{U}} \tag{3}$$

Figure 4 shows the results of spark test under steady cross-flow of air on the spark gap. The maximum values are also shown for each inflow condition (square markers) to

highlight the variation in the measurement volume. The turbulent velocity (calculated from the root mean square of the fluctuating velocity) increases with increasing inflow velocity. The spark plug electrodes act like bluff bodies in the flow, generating turbulence in their wake. The variation in turbulent velocity in the spark gap is limited at low inflow velocities but this variation increases at the highest inflow velocity of 100 m/s. However, given the small dimensions of the electrodes and the gap itself, the turbulent velocity does not increase at the same rate as the mean stream velocity when the inflow velocity increases. Therefore, the turbulence intensity shows a decreasing trend with increasing inflow velocity.

At no flow condition, after the spark breakdown, a steady arc can be maintained between the two electrodes [21]. However, with increase in the cross-flow velocity, the arc stretches until it breaks and must be re-established, which leads to the oscillations in the current signal as shown in Fig. 5, left. The frequency of the re-establishment of this plasma channel is termed as the restrike frequency. The restrike frequency increases with increase in the cross-flow velocity as shown in Fig. 5 right. Figure 6 shows the simulation results for the turbulent velocity (left) and the vorticity (right) in the form of contour plots at three different cross-flow velocities of 10, 40, and 100 m/s. The turbulent velocity predicted downstream of the spark plug is probably due to the bluff body effect. The center of the spark gap typically has no turbulence, but the turbulence around the electrodes may affect the plasma channel. The recirculation zone downstream of the gap increases in area with increasing cross-flow velocity. Again, the range of turbulent velocity also increases with increasing cross-flow velocity which causes the increase in the turbulent velocity in the gap (Fig. 4). The vector quantity of vorticity can be physically interpreted as the number of rotations of the eddies per unit time. Mathematically, vorticity represents the curl of the velocity vector. The tensor form of vorticity calculations is given by Eq. (4) [23].

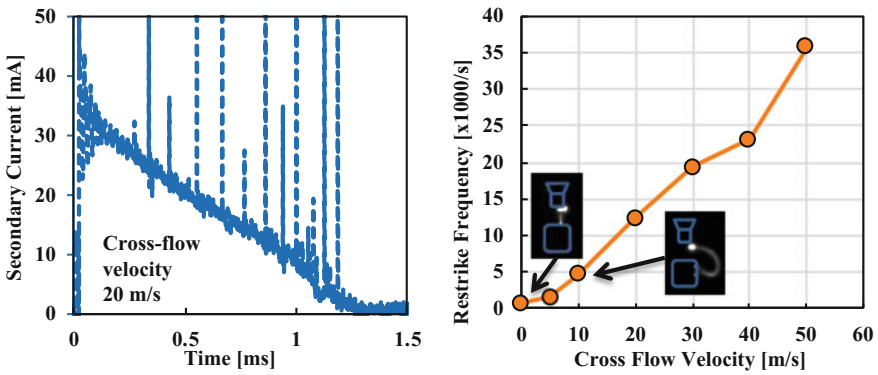


Fig. 5. Spark test – secondary current for 20 m/s cross-flow velocity (left) and restrike frequency (right) [21]

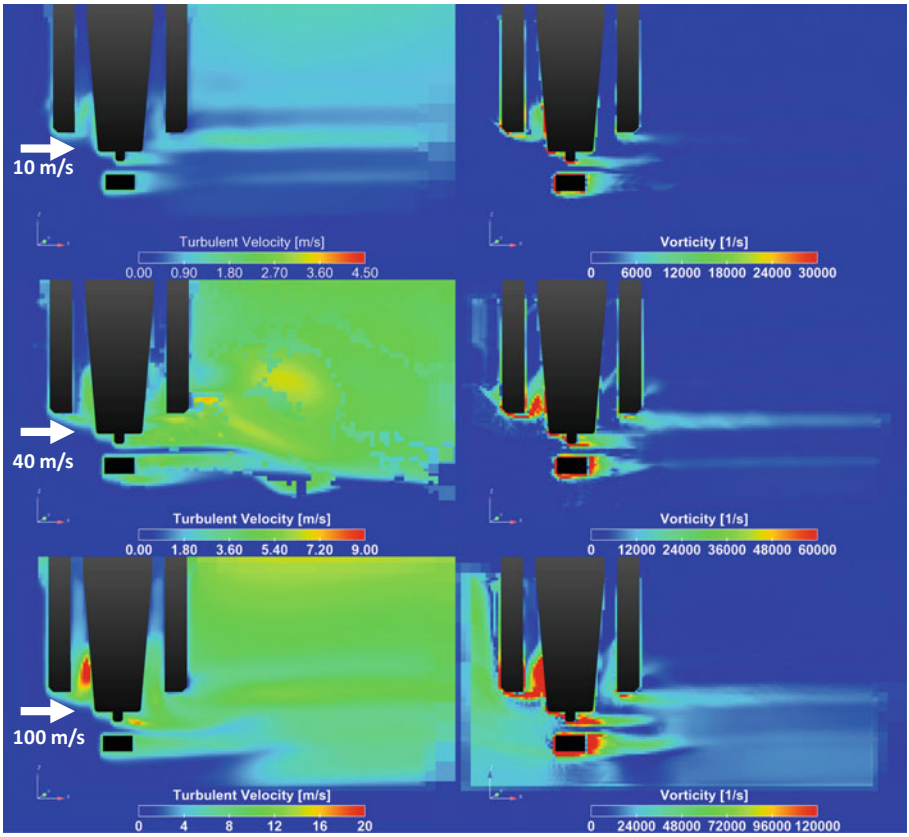


Fig. 6. Steady cross-flow at 10, 40, 100 m/s – turbulent velocity (left) and vorticity (right)

$$\omega = \nabla \times \bar{u} = \left(\frac{\partial u_k}{\partial x_j} - \frac{\partial u_j}{\partial x_k}\right)\hat{\imath} - \left(\frac{\partial u_k}{\partial x_i} - \frac{\partial u_i}{\partial x_k}\right)\hat{\jmath} - \left(\frac{\partial u_j}{\partial x_i} - \frac{\partial u_i}{\partial x_j}\right)\hat{k} \tag{4}$$

where u_i, u_j, and u_k are the x, y, and z components of the velocity, respectively.

As illustrated in the contour plots, the vorticity increases with increasing cross-flow velocity. Vorticity and restrike frequency follow a similar increasing trend with increasing the cross-flow velocity. At 10 m/s and 40 m/s cross-flow velocities, the vorticities are in the range of 6000–12000 s^{-1} and 24000–36000 s^{-1}, which seems to correlate with the restrike frequencies of the spark discharge (Fig. 5) at these cross-flow velocities (5000 s^{-1} and 24000 s^{-1}, respectively). The cross-flow simulation study gives an estimation of the influence of the spark plug geometry in generating turbulence in its wake.

4.2 Engine Motoring Simulation

In this subsection, the results for simulation of one gas exchange cycle of the engine are described. A cycle is described from 0 to 720 °CA with the compression TDC at 360 ° CA. Increasing the mass air flow (MAF) of the engine while maintaining the same fuel injection quantity would make the fuel–air mixture leaner. Two types of methods to increase the MAF into the engine are studied—increasing the throttle opening and increasing intake pressure at a constant throttle opening. The typical spark timing window of 300 to 360 °CA is investigated.

Since the focus of this research is the flow field around the spark plug, two planes are selected for analysis which pass through the electrode gap denoted by the cross-flow plane and the J-plane (illustrated in Fig. 7). Similar to the steady cross-flow analysis, the average of the magnitudes of velocity and root mean square velocity, and the turbulence intensity are calculated in the spark gap over a cylindrical volume which is 0.9 mm in diameter and 0.77 mm in height bringing the total volume to approximately 0.5 mm^3. This gap averaging volume is highlighted in Fig. 7. Similar to the validation conditions, the engine speed and intake temperature are constant.

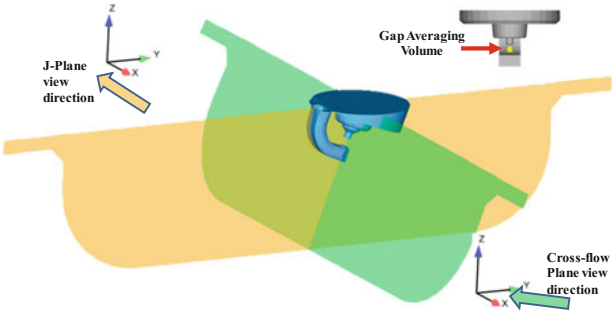


Fig. 7. Result analysis planes and gap averaging volume

4.2.1 Increasing Throttle Opening

During the engine test, the throttle is opened gradually to increase the MAF from 3.05 to 4.8 g/s under normally aspirated intake condition. The corresponding manifold pressure profiles are used as the boundary conditions for the inflow boundary in the simulations. In this MAF range, there is no significant difference in the calculated magnitude of the flow velocity in the spark gap (Fig. 8, left). The overall flow velocity magnitude is highest between 310 to 320 °CA and shows a decreasing trend as the piston approaches TDC.

The turbulence intensity increases since the gap velocity magnitude decreases (plotted with round markers in Fig. 8, right) but the turbulent velocity magnitude remains in the range of $\sim$1 m/s. The turbulent kinetic energy (TKE) decreases since it is a function of the gap velocity (plotted with square markers in Fig. 8, right). Again, the two MAF levels show similar results. Figure 9 shows plots of the maximum velocity magnitudes on each of the analyses planes in the gap. The maximum velocity

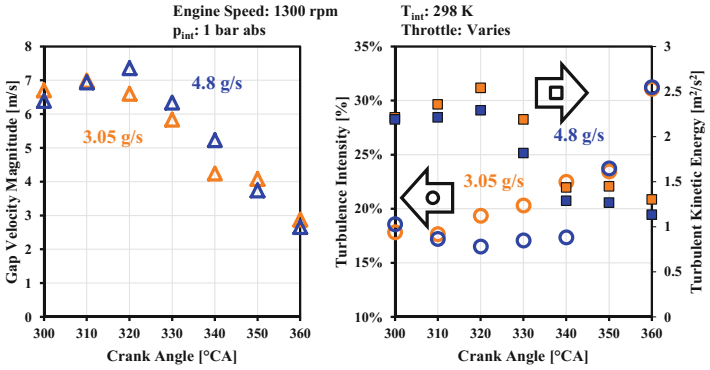


Fig. 8. Throttle opening increase – gap velocity magnitude (left) and turbulence intensity and TKE (right)

in the cross-flow plane is estimated to be between 315 and 330 °CA, and the trends of velocity are similar at the three MAF levels. The velocity in the J-plane reaches a minimum magnitude between 325 and 335 °CA. The velocities in the J-plane for the 4.2 and 4.8 g/s MAF cases are similar, but the velocity is highest in the 3.05 g/s case.

In order to interpret these results, a detailed study of the flow field in the vicinity of the spark plug is required. The velocity and the vorticity contours are shown in Figs. 10 and 11, respectively, in the cross-flow plane over the 310–360 °CA range for the 4.8 g/s MAF case. The velocity vectors are also shown in Fig. 10. Between 310 and 340 °CA, the flow in the gap is directed from −X to +X direction (coordinate system in Fig. 7). Thereafter, the flow direction begins to switch which can explain the local minima in the maximum cross-flow velocity magnitude in the gap (Fig. 9, left). The flow field in the cylinder is affected by both the geometry of the intake port and the geometry of the combustion chamber. As the piston reaches TDC toward the end of the compression stroke, the squish effect can cause a change in the global air motion in the

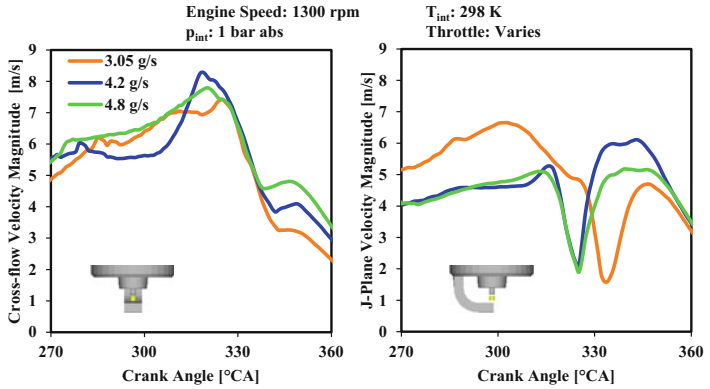


Fig. 9. Maximum velocity magnitude in the analysis planes

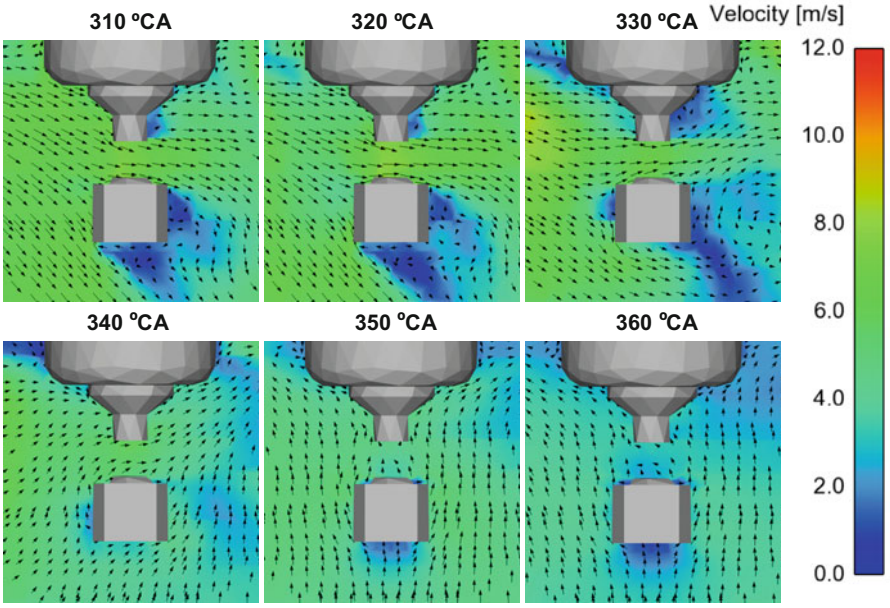


Fig. 10. Velocity contours with vector arrows in the cross-flow plane for MAF of 4.8 g/s

cylinder which in turn can cause the switch in the flow direction through the gap. There is a recirculation zone formed downstream of the ground electrode which is most evident at 330 °CA. The vorticity magnitude in this zone and in the gap is in the range of 6000–8000 s^{-1}, which is similar to the steady cross-flow vorticity results for ~8–10 m/s cross-flow velocity (Fig. 11). The vorticity decreases as the piston approaches TDC. This would imply that some of the best conditions for sustaining the arc would be closer to the TDC after 340 °CA. However, in the case of lean combustion, the spark timing typically has to be advanced to provide sufficient time for the combustion to complete. Therefore, it may be required to initiate and sustain ignition under high flow conditions.

Figure 12 shows the velocity and vorticity contours in the J-plane for the 4.8 g/s MAF case. Analogous to the cross-flow plane, the bulk flow direction changes from outward of the ground electrode to heading toward the ground electrode at 350 °CA. Again, this indicates a switch in the bulk in-cylinder gas motion. There are two recirculation zones in the −Y and the +Y direction (global coordinates in Fig. 7) on both sides of the central electrode with vorticities in the range of 6000–8000 s^{-1}. The vorticity decreases as the piston moves closer to TDC. The ground electrode may play a crucial role in the spark kernel formation. It could shield the gap from flow effects during the initial stage of the spark discharge, and later may hinder the flame propagation as the flame kernel ignites the remainder of the fuel–air mixture. Under this particular operating condition and for this engine geometry and spark plug orientation, it would seem more beneficial for the flow to assist the flame in propagating away from

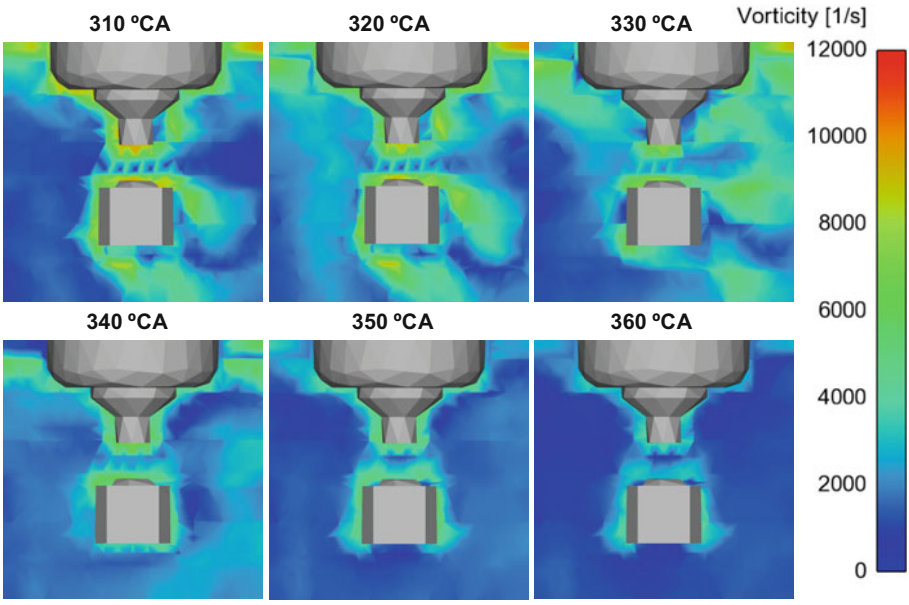


Fig. 11. Vorticity contours in the cross-flow plane for MAF of 4.8 g/s

the ground electrode and into the combustion chamber which may not occur at late spark timings.

4.2.2 Increasing Intake Pressure

The next set of results simulates the engine operating conditions at which the throttle opening is held constant, and the intake pressure is increased to increase the MAF from 7.4 g/s to 15.1 g/s. Again, the manifold pressure profile is used as the pressure boundary condition for the inflow boundary. Figure 13, left shows the magnitude of the gap velocity for the 7.4 and 15.1 g/s MAF cases. The trends differ significantly with piston movement. With increasing intake pressure, the gap velocity decreases. The maximum gap velocity in the analysis planes is given in the appendix.

The turbulent intensity and the turbulent kinetic energy are inversely and directly correlated to the gap velocity, respectively (Fig. 13, right). Beyond 330 °CA, the turbulent kinetic energies for 7.4 and 15.1 MAF cases are similar. The increase in intake pressure is expected to enhance the flow in the cylinder due to a higher pressure differential. However, it may not necessarily increase the flow velocity in the spark gap. In fact, increase in intake pressure and consequent increase in the MAF increase the overall in-cylinder swirl ratio and tumble ratio (shown in Fig. 14). The 3.05 g/s MAF case is also shown for reference. Due to the orientation of the spark plug (Fig. 7), the gap velocity is affected primarily by the swirl motion and the tumble across Y-axis. Flow is shielded by the ground electrode in the X-axis tumble plane. Swirl ratio during the suction stroke increases with an increase in intake MAF, but this swirl motion is not sustained through the compression stroke and the trend is inverted near the spark

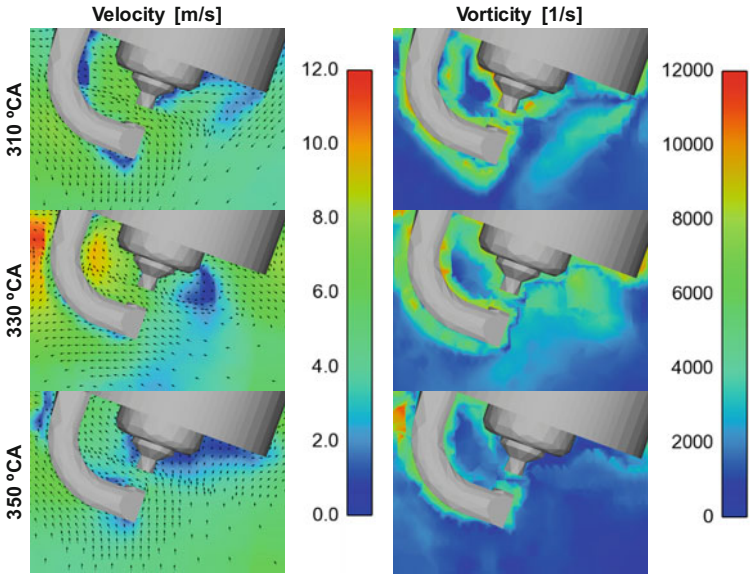


Fig. 12. Velocity and vorticity contours in the J- plane for MAF of 4.8 g/s

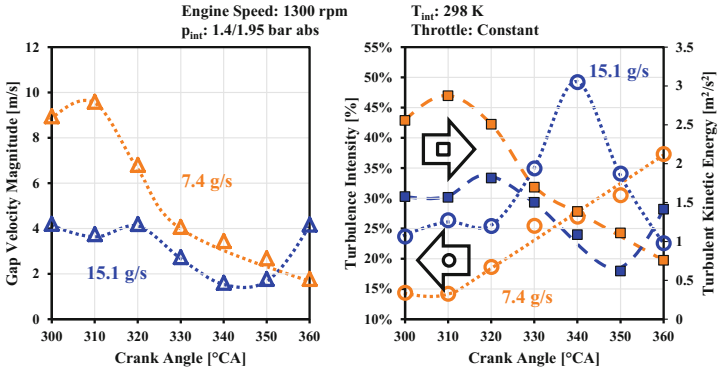


Fig. 13. Intake pressure increase – gap velocity magnitude (left) and turbulence intensity and TKE (right)

timing window. MAF seems to have negligible effect on the tumble around Y-axis near the spark timing window. This reduction in intensity of swirl motion near the spark window can contribute toward reduced flow velocity magnitude in spark gap. Additionally, a reversal in cross-flow direction can be observed near TDC (shown in Fig. 15), which may be a consequence of the change in overall in-cylinder flow field. This implies that the flow in the spark gap can reduce even when the bulk gas velocity is boosted.

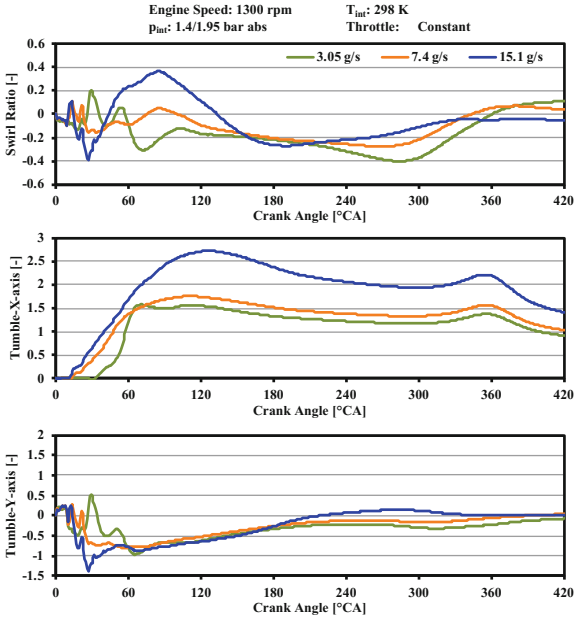


Fig. 14. Intake pressure increase – swirl ratio and tumble ratio across the X- and Y-axes

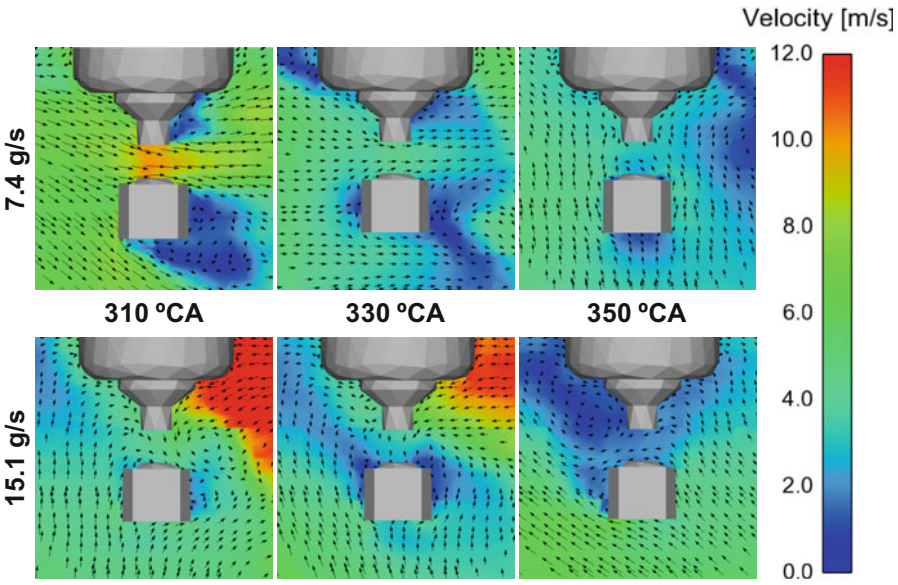


Fig. 15. Velocity contours with vector arrows in the cross-flow plane for MAF of 7.4 and 15.1 g/s

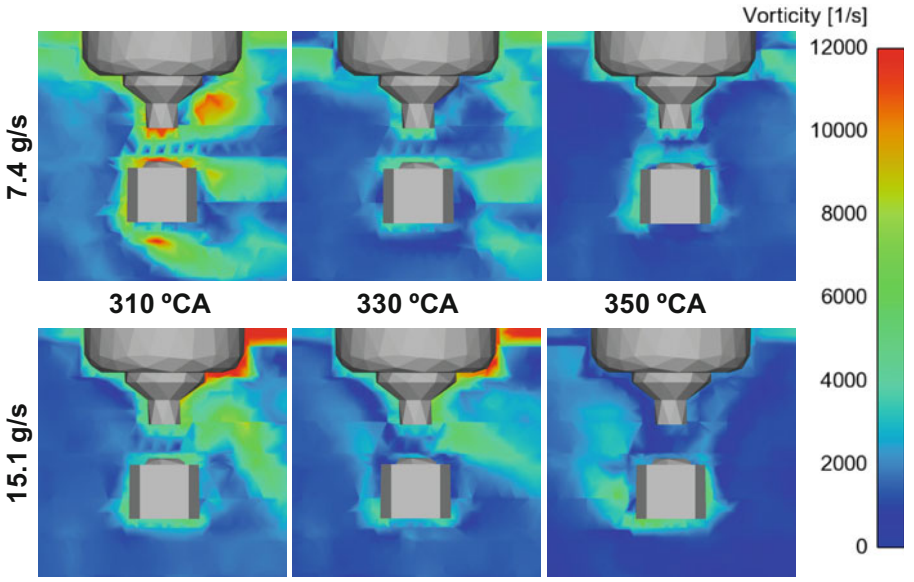


Fig. 16. Vorticity contours in the cross-flow plane for MAF of 7.4 and 15.1 g/s

Figure 15 illustrates a comparison of the velocity contours at the two MAF cases along with the velocity vectors in the cross-flow plane. The flow field around the spark plug on the cross-flow plane is distinct for each case. For the 7.4 g/s case, the flow direction changes beyond 330 °CA possibly due to the change in the flow direction of the bulk gas. The swirl ratio also changes sign at the same crank position (Fig. 14) for the 7.4 g/s case. Moreover, there is the formation of a recirculation zone downstream of the ground electrode. Highest vorticity magnitude in the gap is calculated at 310 °CA. The vorticity in this case is again in the range of 6000–8000 s^{-1} and reduces subsequently as the turbulence decays.

The peak velocity of $\sim$10 m/s is determined at 310 °CA and reduces as the piston moves toward the TDC (Fig. 15). For the 15.1 g/s case, the flow field is much more dynamic. However, the gas flow velocity in the gap never exceeds 5 m/s in the 300 to 360 °CA range. Higher velocity regions (>12 m/s) can be seen near the spark plug which indicates that the bulk gas velocity may be higher than the gap flow velocity. Various recirculation zones are also predicted near the electrodes with vorticity in most of the regions in the range of 6000–8000 s^{-1} (Fig. 16). Maximum vorticity around the gap can reach 8000–10000 s^{-1}. Vorticity reduces as TDC is approached. For this particular engine configuration and operating conditions, with increasing pressure, the flow velocity in the spark gap reduces.

In the J-plane, the vorticity of the recirculation zones around the electrodes is higher for the 15.1 g/s case compared to the 7.4 g/s case (Fig. 17). The vorticity magnitude in the gap on the J-plane for the 15.1 g/s case is lower than the 7.4 g/s case and is in accordance with the reduced flow strength in the spark gap with increasing pressure.

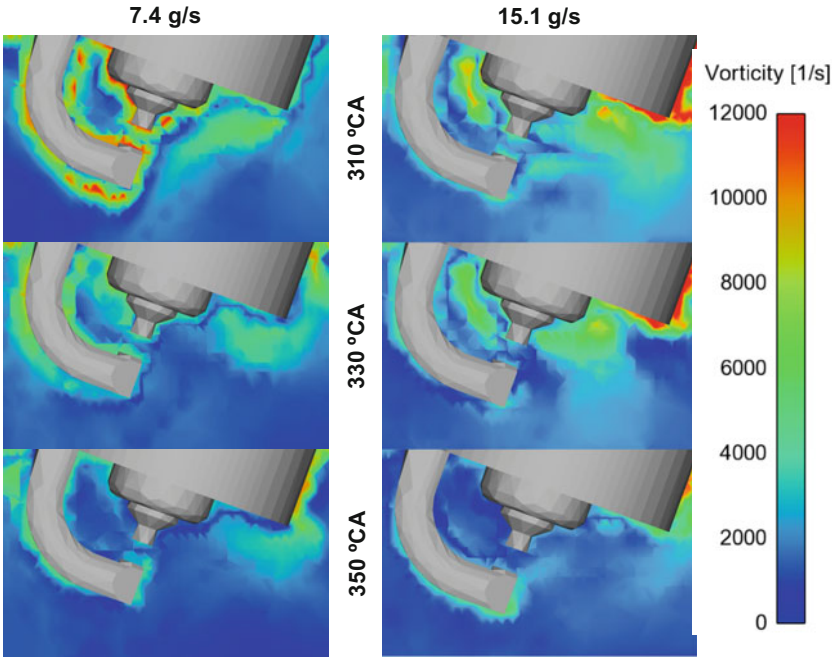


Fig. 17. Vorticity contours in the J-plane for MAF of 7.4 and 15.1 g/s

Detailed empirical studies of the in-cylinder flow field are required to validate the simulation results. Such studies typically require engines with optical access. An indirect method could be analysis of the spark current profile since a relationship between the empirical spark restrike frequency and the simulated vorticity is observed in this study.

5 Summary

The authors performed CFD analysis of the flow fields around a spark plug under steady and uniform flow, and under engine motoring conditions. Motoring simulations were performed with increasing MAF. Two methods to increase the intake MAF were studied—opening throttle and increasing intake pressure. These simulations utilized the empirical intake manifold pressure profile as the inflow boundary condition. The engine motoring simulation cases were validated with engine pressure data to ensure that the gas exchange process was accurately modeled. The main conclusions of this CFD study are as follows:

- Under steady and uniform cross-flow, the spark plug acts like a bluff body, and turbulence is generated in the wake of the plug.

- The vorticity of the eddies which are created in the wake of the spark plug cross-flow is of the same order as the frequencies of the arc restrike in empirical study of the spark discharge with cross-flow.
- There are no significant differences in the flow profiles with increasing throttle opening over the MAF range studied. The gap velocity and turbulent kinetic energy decrease as the piston approaches TDC. The gap velocity direction is also switched near TDC. Different recirculation zones are identified with vorticity magnitudes similar to the steady cross-flow cases.
- When the intake pressure is increased to increase the MAF, there are major differences in the flow field between different MAF cases. The gap velocity decreases with increasing MAF. The bulk gas flow velocity increases with increasing MAF but it may not increase the flow speed in the spark gap.
- The gap velocity and the turbulent kinetic energy in the gap decrease as the piston approaches TDC.
- The vorticity of the flow field around the spark gap is higher for the highest pressure case in comparison to the lowest pressure case.

Acknowledgements. The authors gratefully acknowledge Convergent Science for the use of their simulation suite (version 2.3) and ANSYS for the use of EnSight (version 10.1). Furthermore, the authors acknowledge the Canada Research Chair program, NSERC, CFI, OIT, the University of Windsor, Ford Motor Company, and other OEMs for their support of the research at the Clean Combustion Engine Laboratory.

Nomenclature

AMR	Adaptive Mesh Refinement
BTDC	Before Top Dead Center
CFD	Computational Fluid Dynamics
DC	Direct Current
EGR	Exhaust Gas Recirculation
IC	Internal Combustion
k	Turbulent kinetic energy
MAF	Mass Air Flow
NOx	Nitrogen Oxides
RNG	Renormalized Group
SI	Spark Ignition
TDC	Top Dead center
$\overline{U}$	Mean flow velocity
u’	Turbulent velocity
ε	Dissipation Rate
ω	Vorticity

Appendix: Maximum Velocity Magnitude in the Analysis Planes with Increasing Intake Pressure

See Fig. 18.

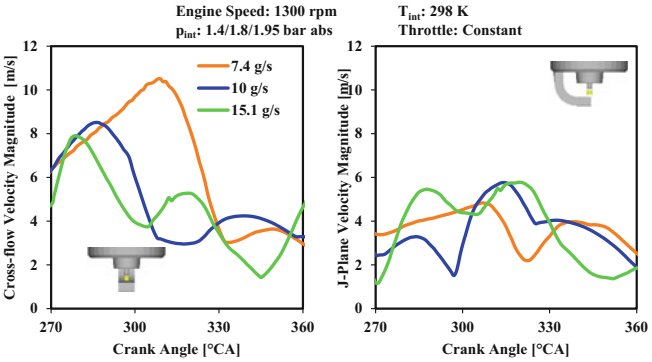


Fig. 18. Maximum velocity magnitude in analysis planes with increasing intake pressure

References

1. National Research Council: Transitions to Alternative Vehicles and Fuels. The National Academies Press, Washington, DC (2013)
2. National Research Council: Review of the Research Program of the FreedomCAR and Fuel Partnership: Third Report. The National Academies Press, Washington, DC (2010)
3. Hsieh, W.-D., Chen, R.-H., Wu, T.-L., Lin, T.-H.: Engine performance and pollutant emission of an SI engine using ethanol–gasoline blended fuels. Atmos. Environ. **36**, 403–410 (2002)
4. Yang, Z., Bandivadekar, A.: 2017 global update: light-duty vehicle greenhouse gas and fuel economy standards. The International Council on Clean Transportation (2017). https://www.theicct.org/publications/2017-global-update-LDV-GHG-FE-standards. Accessed 11 May 2018
5. Germane, G.J., Wood, C.G., Hess, C.C.: Lean combustion in spark-ignited internal combustion engines-a review. SAE Technical Paper 831694 (1983)
6. Attard, W.P., Blaxill, H.: A lean burn gasoline fueled pre-chamber jet ignition combustion system achieving high efficiency and low NOx at part load. SAE Technical Paper 2012-01-1146 (2012)
7. Wei, H., Zhu, T., Shu, G., Tan, L., Wang, Y.: Gasoline engine exhaust gas recirculation–a review. Appl. Energy **99**, 534–544 (2012)
8. Diana, S., Giglio, V., Iorio, B., Police, G.: A strategy to improve the efficiency of stoichiometric spark ignition engines. SAE Technical Paper 961953 (1996)
9. Takahashi, D., Nakata, K., Yoshihara, Y., Ohta, Y., Nishiura, H.: Combustion development to achieve engine thermal efficiency of 40% for hybrid vehicles. SAE Technical Paper 2015-01-1254 (2015)

10. Nakata, K., Nogawa, S., Takahashi, D., Yoshihara, Y., Kumagai, A., Suzuki, T.: Engine technologies for achieving 45% thermal efficiency of S.I. Engine. SAE Int. J. Engines **9**, 179–192 (2016)
11. Kuo, T.-W.: What causes slower flame propagation in the lean-combustion engine? J. Eng. Gas Turbines Power **112**, 348–356 (1990)
12. Arcoumanis, C., Whitelaw, J.H.: Fluid mechanics of internal combustion engines—a review. Proc. Instit. Mech. Eng. Part C: J. Mech. Eng. Sci. **201**, 57–74 (1987)
13. Ives, M.E.: Enhancement of intake generated swirl to improve lean combustion. MASc thesis, Electronic Theses and Dissertations. University of Windsor, Windsor, ON, Canada (2017)
14. Ozdor, N., Dulger, M., Sher, E.: Cyclic variability in spark ignition engines a literature survey. SAE Technical Paper 940987 (1994)
15. Shen, H., Hinze, P.C., Heywood, J.B.: A model for flame initiation and early development in si engine and its application to cycle-to-cycle variations. SAE Technical Paper 942049 (1994)
16. Johansson, B.: Cycle to cycle variations in S.I. Engines-the effects of fluid flow and gas composition in the vicinity of the spark plug on early combustion. SAE Technical Paper 962084 (1996)
17. Sayama, S., Kinoshita, M., Mandokoro, Y., Fuyuto, T.: Spark ignition and early flame development of lean mixtures under high-velocity flow conditions: an experimental study. Int. J. Engine Res. (2018). https://doi.org/10.1177/1468087417748517
18. Aleiferis, P.G., Taylor, A.M.K.P., Whitelaw, J.H., Ishii, K., Urata, Y.: Cyclic variations of initial flame kernel growth in a honda VTEC-E lean-burn spark-ignition engine. SAE Technical Paper 2000-01-1207 (2000)
19. Arcoumanis, C., Bae, C.-S.: Visualization of flow/flame interaction in a constant-volume combustion chamber. SAE Technical Paper 930868 (1993)
20. Ballal, D.R., Lefebvre, A.H.: The influence of spark discharge characteristics on minimum ignition energy in flowing gases. Combust. Flame **24**, 99–108 (1975)
21. Yu, X., Yang, Z., Yu, S., Ives, M., Zheng, M.: Discharge characteristics of current boosted spark events under flow conditions. In: ASME Internal Combustion Engine Division Fall Technical Conference, 15–18 October 2017, pp. V001T03A017 (2017)
22. Schneider, A., Leick, P., Hettinger, A., Rottengruber, H.: Experimental studies on spark stability in an optical combustion vessel under flowing condition. In: Liebl, J., Beidl, C. (eds.) Internationaler Motorenkongress 2016, pp. 327–348. Springer Fachmedien Wiesbaden (2016)
23. ConvergeCFD, ConvergeCFD Manual Series–Converge 2.3 Manual, Convergent Science (2016)

Combustion Characteristics and Emissions of Direct Injection Neat n-Butanol in a Compression Ignition Engine

Zhenyi Yang, Xiaoye Han, Shui Yu, Shouvik Dev, Graham Reader, David S-K Ting, and Ming Zheng(✉)

Department of Mechanical Engineering, University of Windsor, 401 Sunset Avenue, Windsor, ON N9B 3P4, Canada
yangl44@uwindsor.ca

Abstract. Internal combustion engines have been the major power source in the transportation sector for decades. Despite the recent developments in electric vehicles, internal combustion engines will continue to be one of the main propulsion methods for transportation sector, especially in heavy-duty applications. In order to maintain the competitiveness of internal combustion engines, renewable biofuels have been proposed to replace fossil fuels for internal combustion. The application of biofuels can greatly reduce the crisis of fossil fuel shortage and it is beneficial for reducing the life cycle carbon dioxide emissions. In this research, direct injection of neat n-butanol fuel is investigated in a compression ignition research engine. The combustion characteristics of n-butanol differ from that of diesel fuel by the longer ignition delay, shorter combustion duration, higher pressure rise rate, and narrower injection timing window. The NOx and smoke emissions from n-butanol combustion are significantly lower than those of diesel combustion. Different engine loads ranging from 2 bar to 7.5 bar are achieved with n-butanol combustion without the application of EGR. The combustion becomes more intensive with increased load, and the operation timing window is narrower due to the limits imposed by the high-pressure rise rate. To slow down the pressure rise rate, 30% EGR is applied under a load level of 6.5 bar IMEP. It is shown that EGR is not only effective in reducing the pressure rise rate but also beneficial for further reducing the NOx emissions. An EGR sweep experiment is carried out at a higher engine load of 10 bar IMEP with a CA50 of 365 °CA. Though the pressure rise rate decreases with the increase in EGR ratio, a significant increase in smoke emissions is observed when the EGR ratio increases above 35%. The trade-off between smoke emissions and the pressure rise rate can be alleviated by retarding the combustion phasing, but with a penalty of reduced indicated thermal efficiency.

A. Vasel and D. S.-K. Ting (Eds.): EAS 2018, SPE, pp. 55–79, 2019.
https://doi.org/10.1007/978-3-030-00105-6_4

Nomenclature

CA	Crank angle	HFID	Heated flame ionization detector
CA5	Crank angle of 5% total heat release	HRR	Heat release rate
CA50	Crank angle of 50% total heat release	IMEP	Indicated mean effective pressure
CA95	Crank angle of 95% total heat release	NDIR	Non-dispersive infrared detector
CO	Carbon monoxide	NOx	Nitrogen oxides
COV	Coefficient of variation	ppm	Parts per million
CO2	Carbon dioxide	p_inj	Pressure of injection
DI	Direct injection	SOI	Start of injection
EGR	Exhaust gas recirculation	TDC	Top dead center
FSN	Filter smoke number	THC	Total hydrocarbons
HCLD	Heated chemiluminescence detector	dp/dθ_max	Maximum pressure rise rate
HCCI	Homogeneous charge compression ignition		

1 Introduction

The energy consumption has been continuously increasing with the development of the society. According to the Annual Energy Outlook 2018 [1] from the U.S. Energy Information Administration, a steady worldwide increase in energy consumption will continue throughout the projection year of 2050 (Fig. 1). The transportation sector as one of the major energy consuming sectors accounts for about 30% of the total energy consumption. Correspondingly, it is the main contributor of carbon dioxide emission (Fig. 2). The requirements on increased energy demands and decreased production of greenhouse gas and other exhaust emissions raise great challenges for the transportation sector. Electrical vehicles are considered promising as they are efficient and produce nearly zero emissions from the power train; however, the significantly lower energy density of battery compared to liquid fuel is still a major unsolved obstacle for electrical vehicles, especially in heavy-duty applications. In the foreseeable future, internal combustion engines will still be one of the major power sources in the transportation sector. In order to maintain the competence of internal combustion engines, various techniques have been developed to improve the engine efficiency and reduce engine-out emissions. In addition, the utilization of renewable energy sources in internal combustion engine can potentially decrease life cycle greenhouse gas emissions and contribute to worldwide energy security [2].

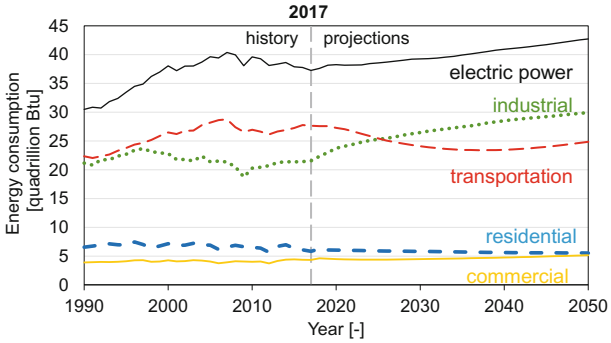


Fig. 1 Energy consumption by sector (Reference case, adapted from [1])

The application of biofuels in internal combustion engines has been studied for decades [3]. Previous research results have proved that the modification of fuel properties by using alternative fuels is effective in controlling the in-cylinder combustion process and reducing engine-out emissions [4–6]. Alternative fuels can be applied in internal combustion engines through different ways. They can be used in blended fuels [7–16], as a neat fuel [17–26], or mixed with conventional fuel via a secondary injector for the dual-fuel combustion [27–30]. Commonly used alternative fuels include butanol and ethanol. Butanol and ethanol are both renewable alternative fuels that can be produced from biomass fermentation. There are several different butanol isomers: n-butanol, sec-butanol, iso-butanol, and tert-butanol. In this research, n-butanol is used and is referred to as butanol hereafter for simplification. The properties of n-butanol and ethanol are compared with diesel fuel as listed in Table 1. Butanol and ethanol are advantageous in reducing the smoke emissions due to the oxygen content in their molecules and their better volatility compared to diesel [4, 19]. Lower NOx emissions are also reported by some researchers [18, 22]. Due to the lower fuel reactivity compared with that of diesel, the combustion of alcohol fuels usually produces higher incomplete combustion products, i.e., HC and CO emissions [19–21, 23]. The

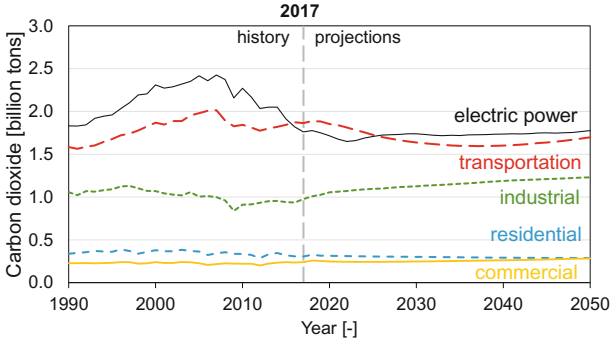


Fig. 2 Energy-related carbon dioxide emissions by sector (Reference case, adapted from [1])

difficulties of applying these fuels mainly come from two aspects: fuel handling and changed combustion characteristics. Butanol has some advantages over ethanol with respect to fuel handling. It is more stable and resistive to water absorption so that it can be transported with existing fuel transportation infrastructures. The lubricity of butanol is better than ethanol. With some lubricity improver, butanol can be used as a direct injection fuel for compression ignition engines [19, 26]. This adds to the attraction of butanol, as the direct injection provides an effective way of controlling the combustion process in the neat fuel application.

Table 1 Comparison of fuel properties

Fuel	Diesel	N-Butanol	Ethanol
Formula	$C_1H_{1.87}$	C_4H_9OH	C_2H_5OH
Density [40 °C, kg/m^3]	0.85	0.81	0.788
Cetane number [−]	44	~25	8~11
Octane number [−]	~25	~27	110~115
Lower heating value [MJ/kg]	45.3	33.1	27
Oxygen content by mass [%]	0	21.62	34.78
Kinematic viscosity @ 40 °C [cSt]	1.9	2.6@25 °C	1.52

The butanol has been applied via different methods in internal combustion engines. As a blended fuel, butanol and diesel with different volume ratios were applied in a compression ignition engine by Rakopoulos et al. [9]. They reported significantly reduced smoke emissions and slightly reduced NOx and CO emissions with the blended fuel. The reduction was more significant when the butanol ratio was high [9]. Iannuzz and Valentino used 20% of butanol mixed with diesel and investigated the combustion and emission performance of the fuel blends. They found that the addition of butanol was in favor of reducing NOx and smoke emission simultaneously [15]. The better performance in reducing smoke emissions was reported by several other researchers using blended diesel and butanol as well [8, 10, 12, 15, 16]. Choi et al. [10] found that there was an upper limit of butanol ratio in the fuel blends as precursors of PAHs might be produced when the blended ratio was too high. The limit of butanol ratio in the blended fuel placed a restriction on the possible benefit in reducing smoke emission with butanol fuel.

The dual-fuel combustion mode with port injected butanol and direct injected diesel was investigated as well [4, 29]. The ratio of butanol and diesel was adjusted dynamically during the test. The general trend for this type of combustion was that the increase in butanol ratio increased the ignition delay and was beneficial for reducing smoke emissions. However, the high-pressure rise rate with increased butanol ratio was undesired and EGR was required to suppress the high-pressure rise rate. This again sets an upper limit for the butanol ratio as a higher butanol ratio required a higher EGR ratio for pressure rise rate reduction, while a higher EGR ratio increased the smoke emissions. The incomplete combustion products in the dual-fuel mode were reported to be higher compared to those of pure diesel combustion [4]. Apart from the change in combustion characteristics of the dual-fuel combustion, the fuel handling was also more complicated as two independent fueling systems were required.

In addition to blended fuel and dual-fuel modes, some researchers used butanol as a neat fuel to fully replace diesel in compression ignition engines. In the neat fuel applications, butanol was either injected via a port fuel injector [4, 22, 25] or injected directly into the cylinder through the high-pressure injection system [19, 26]. For port injected butanol combustion, challenges were reported on properly controlling the combustion process as the combustion was largely in the homogeneous charge compression ignition (HCCI) mode [4, 22]. The unstable combustion at lower load and the high-pressure rise rate at higher load were reported, despite the benefit in reducing NOx and smoke emissions [4, 22, 25]. Han et al. achieved a load of around 9.5 bar IMEP using port fueled butanol when heavy EGR (50%–57%) was applied to suppress the pressure rise rate [4]. The authors also pointed out the control difficulties due to the high sensitivity of the combustion to the changes in gas compositions [4]. Butanol was also used in the direct injection system by adding some lubricity improver [19, 27]. The direct injection provides additional leverage for controlling the combustion process. Zheng et al. used neat butanol in a high compression ratio diesel engine [19]. The injection was conducted via the original diesel common rail injection system. It was seen that butanol produced ultralow NOx and soot emissions [19]. Nevertheless, the injection timing was limited to a narrow window due to the very high-pressure rise rate at early injection timings or unstable combustion at later injection timings [19]. Researchers then tried using different injection strategies to extend the engine load level [19, 26, 27]. Han et al. used neat n-butanol multi-shot injection with EGR dilution and achieved engine load level as high as 19 bar IMEP [27].

Previous research results indicate that butanol is an alternative fuel with great potential to fully replace diesel fuel in compression ignition engines. The combustion characteristics of butanol are, however, significantly different from conventional diesel combustion. These differences result in challenges in the application. Sophisticated control techniques are required to ensure smooth and stable combustion process. Direct injection of butanol is considered a good way as the direct injection greatly increases the control flexibility. Previous research has proved that a very high load can be achieved with butanol using the multi-shot direct injection strategy [27]. The objective of this paper is not to discuss the pathways to increase the engine load levels with butanol fuel, but rather concentrate on the combustion characteristics of neat butanol with direct injection. So, the focus will be on single-shot injection cases. The objective is to investigate the combustion and emissions characteristics of directly injected butanol under different engine operating conditions. The empirical study should contribute to the design of combustion control strategies to accommodate the changed fuel properties, so that high efficiency and clean combustion can be realized with butanol as a renewable biofuel for advanced internal combustion engines. The paper is arranged as follows: first, the combustion characteristics and exhaust emissions of butanol and diesel are compared under similar engine operation conditions; then the butanol combustion under different engine loads is investigated; based on the results from the previous sections, EGR is used to improve the control of the combustion process; the last section presents a study on the effectiveness of EGR with increased engine load.

2 Research Methodology

2.1 Experimental Setup

Engine tests are conducted on a single-cylinder research engine which is converted from a four-cylinder production diesel engine. The configuration of the system is shown in Fig. 3. The specifications of the research cylinder are listed in Table 2. The research cylinder is isolated from other three cylinders and connected with independent intake and exhaust systems. An oil-free compressor is connected to the intake system. The intake pressure is controlled by a pressure regulator. A ROOTS® air flow meter is used to measure the mass flow rate of the intake air. The exhaust backpressure is controlled via an electrically controlled pneumatic backpressure valve. An EGR loop with an EGR cooler and EGR valve is routed from the exhaust pipe to the intake pipe. The EGR ratio is controlled by the adjustment of the backpressure and the opening of the EGR valve.

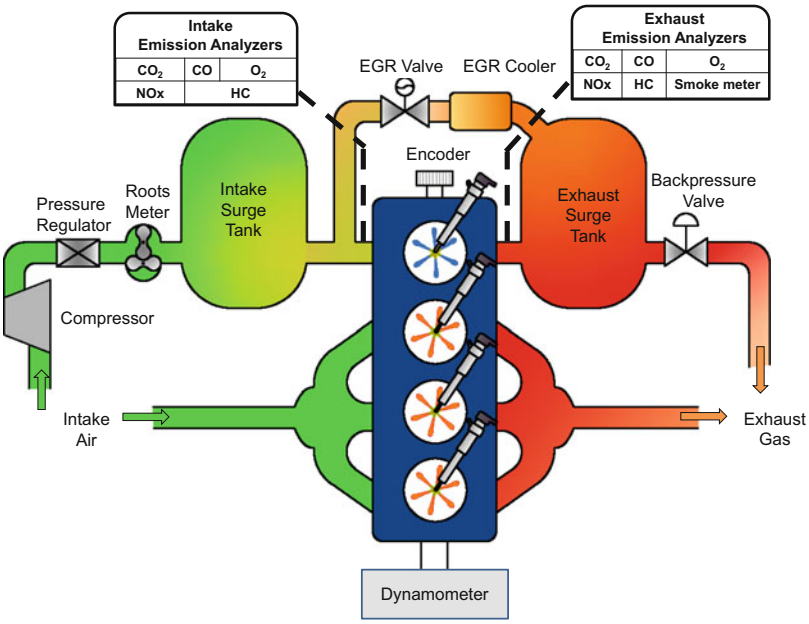


Fig. 3 Schematic of the research engine system

Table 2 Research cylinder specifications

Displacement volume (liters)	0.499
Bore (mm) × Stroke (mm)	86 × 86
Compression ratio (-)	18.2:1
Fuel injection	Common rail direct injection
Injection pressure (bar)	Up to 1600

The fueling system of the engine test platform is composed of two lines. The three non-research cylinders are fueled with conventional diesel, while the research cylinder is equipped with an independent butanol fuel supply line. This fuel supply line consists of a low-pressure feed pump, which feeds fuel into a high-pressure pump. The high-pressure pump increases the fuel pressure and delivers the fuel into a high-pressure common rail. A high-pressure injector is connected to the common rail for the direct fuel injection. In order to protect the high-pressure fuel injection system, 500 ppm of lubricity improver is added into the butanol fuel. The effects of this lubricity improver on the combustion process are not considered in the analysis. The injector is driven by an EFS IPod high-pressure injector driver. The command of the injection is controlled by an in-house developed control system based on real-time (RT) computer embedded with field-programmable gate arrays (FPGA). The fuel injection pressure, injection duration, and injection timing can all be controlled by the fuel control system.

The gas compositions are measured by two sets of emission analyzers. One set of the analyzers is connected to the intake pipe after the junction of fresh air and recirculated exhaust gas, so that the composition of the intake charge flowing into the cylinder is measured. The sampling port is located about 0.5 m downstream of the joint location so that the gases are properly mixed when they reach the measurement location. The exhaust sampling line is connected downstream of the exhaust port. The gas analyzers include a nondispersive infrared detector (NDIR), a paramagnetic detector, a heated chemiluminescent detector (HCLD), and a flame ionization detector (FID). They are used to detect CO, CO_2, O_2, NOx, and HC in the intake and exhaust flow, respectively. An AVL 415S smoke meter is connected to the exhaust sampling line to measure the smoke emissions.

An eddy current dynamometer is connected to the engine for engine torque/speed control and measurement. In addition, in order to maintain a stable engine operation condition, the temperatures of engine oil and coolant are controlled by FEV lubricant and coolant conditioning units, respectively, which can maintain the temperature of the lubricant and coolant during engine operation at a set value of 80 °C.

2.2 Combustion Analysis

In this study, the engine load is defined as

$$\mathrm{IMEP} = \int_0^{720} \frac{\mathrm{p}(\theta)\mathrm{dV}}{\mathrm{V_d}} \tag{1}$$

where IMEP is indicated mean effective pressure, $\mathrm{p}(\theta)$ is the in-cylinder pressure at each crank angle, dV is the volume change between two consecutive crank angles, and $\mathrm{V_d}$ is the displacement volume of the cylinder.

The combustion analysis is based on the measured in-cylinder pressure. The heat release rate is calculated through Eq. (2):

$$\mathrm{HRR} = \frac{1}{\gamma - 1} \times \left[\gamma p \frac{dV}{d\theta} + V \frac{dp}{d\theta} \right] \tag{2}$$

where HRR is the heat release rate, γ is the ratio of specific heats for the cylinder charge, p is the measured cylinder pressure at the specific crank angle, V is the cylinder volume at the specific crank angle, and θ is the crank angle.

The total heat release can be calculated by the accumulation of the heat release rate. The start of combustion is defined as the crank angle when 5% of the total heat is released. The ignition delay is defined as the time duration between the start of fuel injection and the start of combustion. The combustion duration is defined as the time period between the crank angle when 5% of the heat is released and the crank angle when 95% of the heat is released. The combustion phasing is defined as the crank angle when 50% of the heat is released.

The EGR ratio is defined as the ratio between the intake CO_2 concentration and the exhaust CO_2 concentration:

$$\mathrm{EGR} = \frac{[CO_2]_{\mathrm{int}}}{[CO_2]_{\mathrm{exh}}} \tag{3}$$

3 Results and Discussion

3.1 Comparison Between Diesel and Butanol Combustion

Two sets of tests are carried out with diesel and butanol under similar conditions to investigate the change in combustion characteristics caused by the different fuel properties. The injection pressure, boost pressure, and intake oxygen concentration are kept constant in these tests, as listed in Table 3. In order to achieve the same nominal engine load, the fuel mass of butanol is higher than that of diesel due to the lower energy content of butanol fuel.

Table 3 Engine operating conditions for comparison tests between n-butanol and diesel

Nominal IMEP (bar)	6.5
Injection pressure (bar)	600
Boost pressure (bar abs)	2
Intake O_2 (vol. %)	20.7
Engine speed (rpm)	1500
Fuel mass (mg/cycle)	Butanol: 22; Diesel: 16
Start of injection	Sweep

The ignition delays of the two sets of tests are compared in Fig. 4. It can be seen that the ignition delays of butanol are significantly longer than those of diesel. Overall, the ignition delays are less than 2 ms throughout the SOI sweep for diesel combustion.

While in the butanol combustion cases, the ignition delays range from 2 to 4 ms. The in-cylinder temperature and pressure are higher as the piston moves closer to TDC, and thus the injected fuel experiences shorter ignition delay in a hotter environment as the injection timing moves toward TDC. Under the tested cylinder pressure and temperature conditions, butanol exhibits a higher ignition delay sensitivity to the change in injection timing.

Two cases with injection timing of 338 °CA and 347 °CA are marked with solid dots in Fig. 4. Their in-cylinder pressure and heat release rate profiles are compared in Fig. 5. The diesel combustion starts much earlier than the butanol combustion in both injection timing cases. The diesel combustion has a two-stage heat release shape corresponding to the premixed and diffusion combustion as the injection has not finished at the time of ignition. The amplitude of the first stage heat release is higher when the injection timing is at 338 °CA. It is due to the slightly longer ignition delay at this timing as can be seen in Fig. 4. Longer ignition delay results in better mixing so that the premixed portion increases. The butanol combustion exhibits a single-stage heat release shape at both injection timings. The long ignition delay in butanol combustion provides sufficient time for the fuel–air mixing so that the combustion is predominately in the premixed combustion mode. The sweeps of injection timing in the two sets of tests are not within the same time window due to different constraints in each of the test sets, as will be discussed in the following sections.

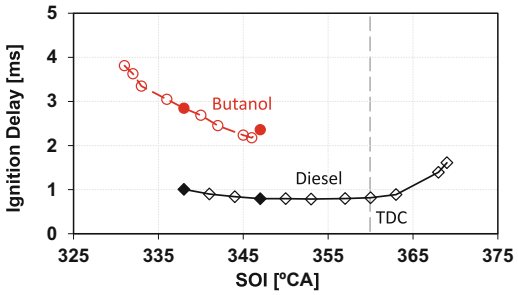


Fig. 4 Butanol vs. diesel—ignition delay

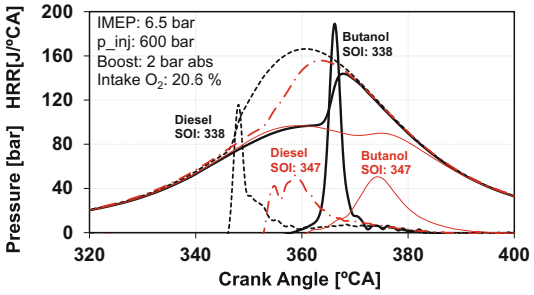


Fig. 5 Butanol vs. diesel—in-cylinder pressure and heat release rate

The combustion phasing is illustrated in Fig. 6. With a short ignition delay of diesel fuel, the combustion phasing is almost linearly correlated to the injection timing. The latest injection timing is limited by misfire where the combustion occurs in the expansion stroke. The earliest injection timing is limited by the high-pressure rise rate when the major part of the combustion process happens during the compression stroke. For butanol combustion, the combustion phasing is distributed in a much narrower window from about 365 °CA to 375 °CA, beyond which the combustion is unstable. The combustion phasing is advanced when the injection timing is moved from 347 ° CA to 336 °CA. Further advancement of the injection timing results in retarded combustion phasing due to the much longer ignition delay. The in-cylinder pressure and heat release rate profiles of butanol combustion in Fig. 7 correspond to the solid dots in Fig. 6.

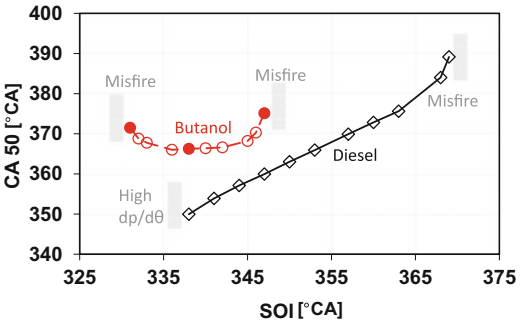


Fig. 6 Butanol vs. diesel—combustion phasing

A stable operation window of butanol combustion with injection timing between 333 °CA and 345 °CA is identified in this condition as can be seen in the top plot of Fig. 8. Further advancement or retardation of the injection timing results in unstable combustion. The combustion durations of the two cases are compared in the bottom plot of Fig. 8. The combustion durations of butanol are shorter compared to that of the diesel fuel. The combustion durations are maintained at around 10 °CA in the stable operation window for butanol. Further advancing or retarding the injection timing results in a longer combustion duration, which ultimately leads to unstable combustion. The stable operation window also correlates to a window with high-pressure rise rate and high indicated thermal efficiency in the butanol combustion regime. The indicated thermal efficiency of butanol combustion can reach about 40%, which is comparable to that of diesel. However, this high-efficiency window is narrower than that of the diesel, and the pressure rise rate of butanol in this window is higher (Fig. 9).

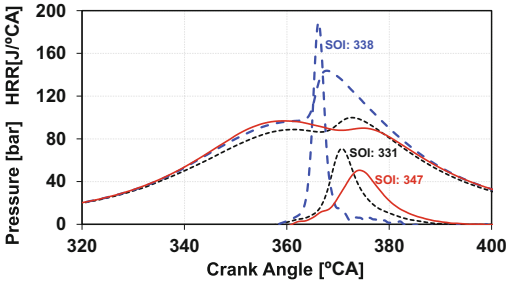


Fig. 7 In-cylinder pressure and heat release rate of butanol combustion at different injection timings

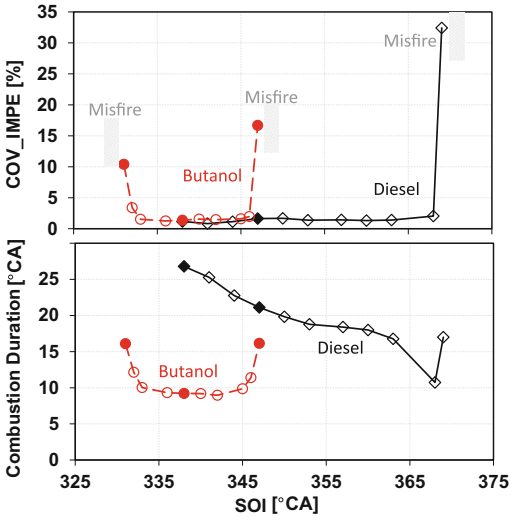


Fig. 8 Butanol vs. diesel—coefficient of variation (COV) of IMEP and combustion duration

The emissions of butanol and diesel combustion are compared in Figs. 10 and 11. Butanol combustion produces significantly lower NOx emissions due to the lean burn of a better mixed mixture. The exhaust smoke emissions are close to zero. The CO and THC emissions are much higher in butanol combustion, even in the stable operation window. It is even worse when the combustion becomes unstable.

According to the test results, butanol exhibits distinct combustion characteristics compared to those of diesel fuel. Due to the drastic differences in fuel reactivity and volatility, the mixing and combustion processes differ between butanol and diesel, and as a result, different fuel injection timing windows are required to ensure stable operation of each fuel. Under the tested conditions, an injection timing window between 333 °CA and 345 °CA is considered suitable for butanol. The pressure and temperature at excessively advanced or retarded injection timings are not appropriate

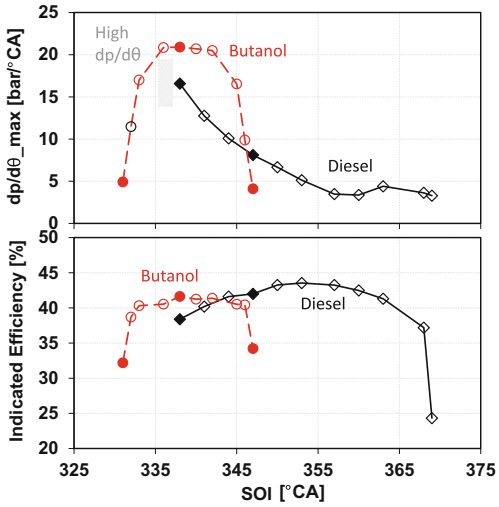


Fig. 9 Butanol vs. diesel—peak pressure rise rate and indicated efficiency

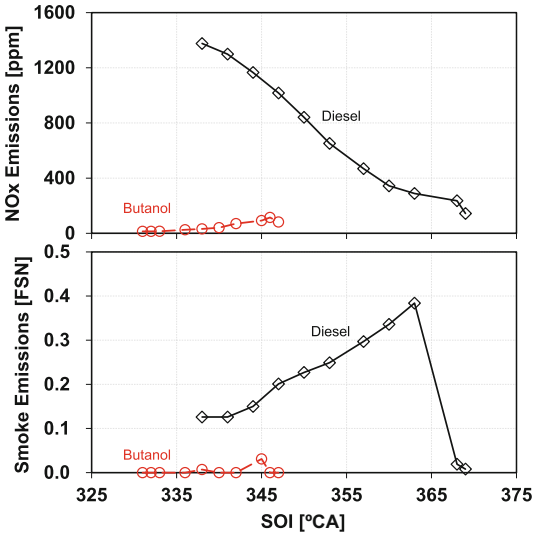


Fig. 10 Butanol vs. diesel—NOx and smoke emissions

for the butanol combustion as can be seen from the high variations of IMEP in Fig. 8. Furthermore, owing to the low reactivity and high volatility, butanol fuel has much longer ignition delay time compared with diesel fuel, which results in better mixing of butanol. This is beneficial for reducing NOx and smoke emissions. However, when the

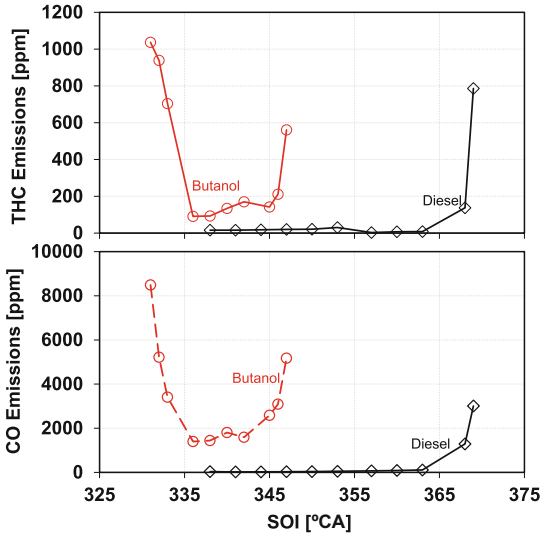


Fig. 11 Butanol vs. diesel—CO and THC emissions

well-mixed air–fuel mixture reaches the combustion conditions and starts to ignite simultaneously, the pressure rise rate is very high, which can result in problems of increased engine noise and excessive mechanical stresses.

3.2 Butanol Combustion Under Different Load Levels

In this section, the combustion performances of butanol under different load levels are investigated and the test conditions are listed in Table 4.

Table 4 Engine operating conditions for butanol combustion under different engine load levels

Nominal IMEP (bar)	2–7.5
Injection pressure (bar)	600
Boost pressure (bar abs)	2
Intake O_2 (vol. %)	20.7
Engine speed (rpm)	1500
Start of injection	Sweep

As can be seen in Fig. 12, the ignition delay decreases with increased engine load under the same injection timing, which can be attributed to the increased fuel amount thus richer mixture, and the higher cylinder temperature under higher load level. The ignition delays are in the range of 2 to 4 ms depending on the injection timing. The pressure rise rate is strongly dependent on the engine load. With the increase in engine load, the peak pressure rise rate increases under the same injection timing. When the

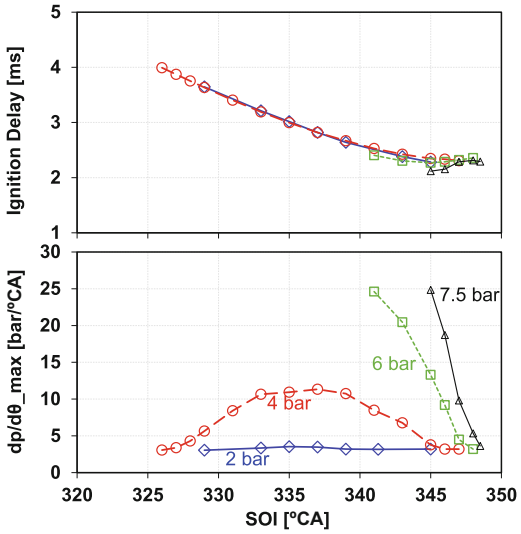


Fig. 12 Ignition delay and peak pressure rise rate at different loads

load level is higher, the peak pressure rise rate increases faster with the advanced injection timing as can be seen from the steeper slope of pressure rise rate in Fig. 12. The peak pressure rise rate reaches 25 bar/°CA for engine load of 6 bar and 7.5 bar cases. From the discussion in the first section, a lower pressure rise rate can also be achieved by further advancement of the injection timing. However, wall impingement may occur when the injection timing is too early in the compression stroke, and therefore it is not attempted. The combustion phasing and stability are shown in Fig. 13. The change of combustion phasing follows the same trend under all the load levels, where it first advances with the advanced injection timing, and then a reverse trend is observed when the injection timing is further advanced. The combustion phasing is slightly earlier under a higher load when the injection timing is the same, which can be attributed to both a shorter ignition delay and a shorter combustion duration. The combustion is highly unstable when the engine load is 2 bar. The COV of IMEP is higher than 10% throughout the SOI sweep. When the load level increases to 4 bar, the operation is more stable, and the stable operation window is about 15 °CA from 329 °CA to 345 °CA. The operation windows of higher load conditions are largely confined by the high-pressure rise rate at early injection timings and the unstable combustion at late injection timings.

The pressure and heat release rate at different load levels with an injection timing of 345 °CA is shown in Fig. 14. The combustion intensity gradually decreased with the decrease in load level as can be seen from the lower heat release profiles. The combustion phasing is retarded, and the combustion duration is longer when the load is decreased. This can be attributed to two main reasons as discussed previously: the lower temperature and the leaner mixture. The equivalence ratios under different engine load levels are shown in Fig. 15. The overall equivalence ratio is reduced with the decrease of load levels so that the reaction is slower due to the lean operation. In order

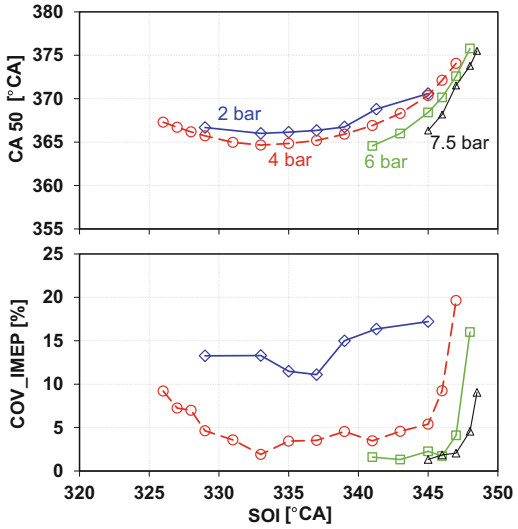


Fig. 13 Combustion phasing and coefficient of variation (COV) of IMEP at different loads

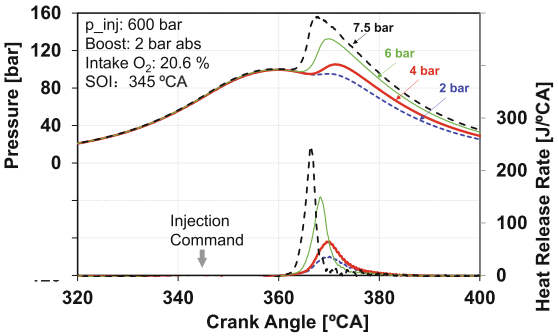


Fig. 14 In-cylinder pressure and heat release rate at different loads

to promote the combustion under lower engine load, various approaches can be used such as making the mixture richer by using lower boost pressure. However, when a lower boost is used, the ignition is problematic which results in unstable combustion. Some other approaches should be added to create a more favorable environment for the combustion, such as intake heating, which is not discussed in this paper. For combustion under higher load, the main concern is the high-pressure rise rate, which creates rough combustion and high noise level. The reason for the high-pressure rise rate is the excessive combustion rate. Techniques which can slow down the combustion rate are beneficial for controlling the pressure rise rate, as reported in the following sections.

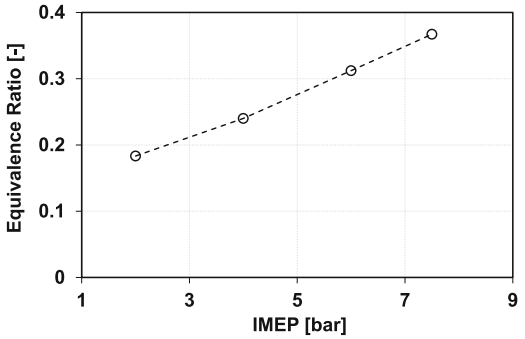


Fig. 15 The equivalence ratio at different loads

The exhaust emissions are shown in Figs. 16 and 17. The NOx emissions are very low under lower engine load and gradually increases with engine load. The smoke emissions are ultralow for all the load levels. The THC and CO emissions increase with decreased engine load, which is resulted from the incomplete combustion.

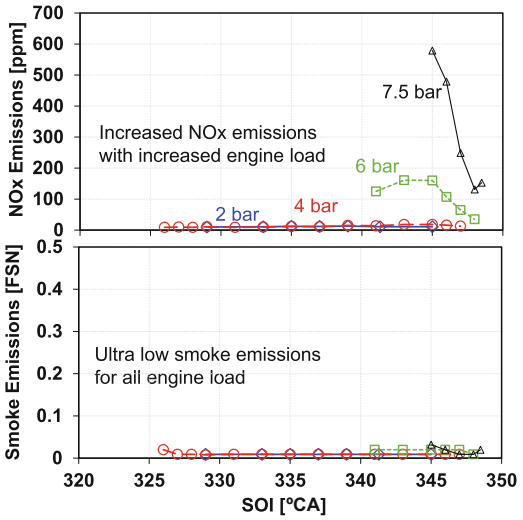


Fig. 16 NOx and smoke emissions at different loads

From the results of injection timing sweep under different engine loads, it can be concluded that butanol has superior performance in reducing NOx and soot emissions compared to diesel; however, the unstable combustion under low-load conditions and the high-pressure rise rate with increased load are challenging for applying butanol in compression ignition engines. The control over the injection timing alone seems insufficient to tackle these problems.

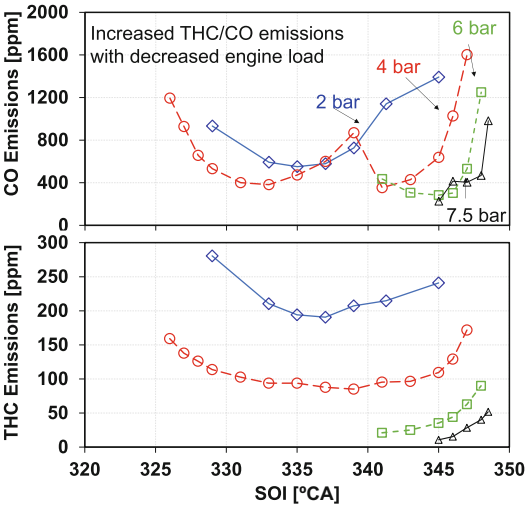


Fig. 17 HC and CO emissions at different loads

3.3 Butanol Combustion with EGR Dilution

The high-pressure rise rate with increased engine load largely limited the operation window of the butanol combustion. In order to suppress the high-pressure rise rate, EGR dilution is applied to slow down the fast burning process. As shown in Fig. 18, with the same injection timing, the combustion initiates later and the combustion duration is longer when 30% EGR is used. Correspondingly, the intensity of the combustion is reduced due to the suppressed reactivity by the addition of EGR. An injection timing sweep is carried out in comparison to the 0% EGR case. The effect of EGR on the combustion duration is shown in Fig. 19. The combustion duration is generally prolonged in the 30% EGR case, with exceptions of very early injection timing where the combustion in the 0% EGR case is also unstable. Due to the longer combustion duration, the peak pressure rise rate is reduced as can be seen in Fig. 20.

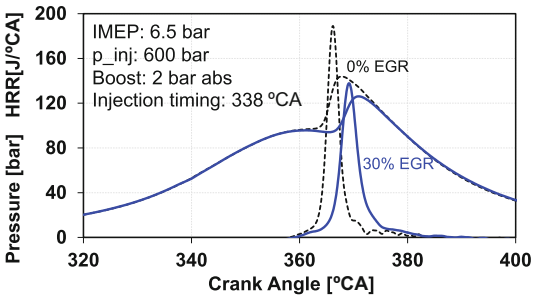


Fig. 18 0% EGR vs. 30% EGR: in-cylinder pressure and heat release rate

The stable operation window is narrowed with the addition of EGR; however, it can maintain the high indicated thermal efficiency in this operation window, as shown in Fig. 21.

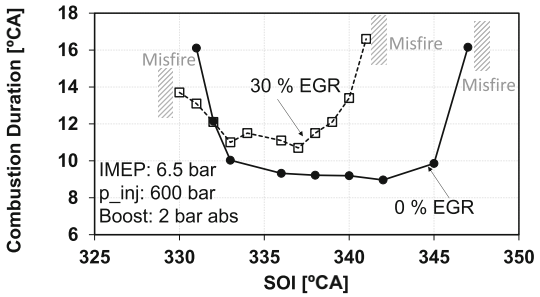


Fig. 19 0% EGR vs. 30% EGR: combustion duration

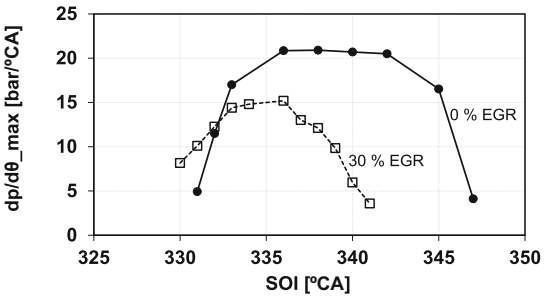


Fig. 20 0% EGR vs. 30% EGR: pressure rise rate

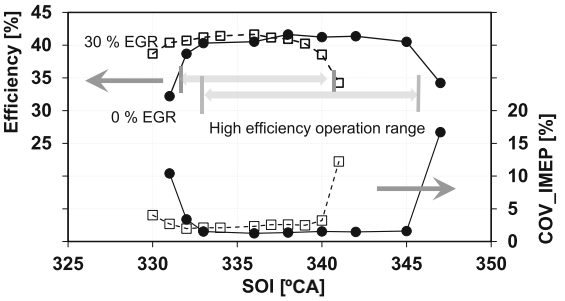


Fig. 21 0% EGR vs. 30% EGR: indicated efficiency

The NOx emissions are further reduced with the application of EGR as can be seen in Fig. 22. The NOx emission under 0% EGR at early injection timing window from 331 to 333 °CA is lower than the NOx emissions with 30% EGR in the same injection

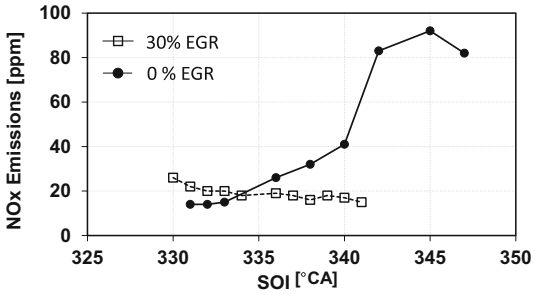


Fig. 22 0% EGR vs. 30% EGR: NOx emissions

timing window. This is because of the unstable combustion in the 0% EGR cases within this time window, as can be seen from the high COV of IMEP in Fig. 21. Meanwhile, the smoke emissions are kept at ultralow levels for both EGR ratios throughout the injection timing sweep (Fig. 23).

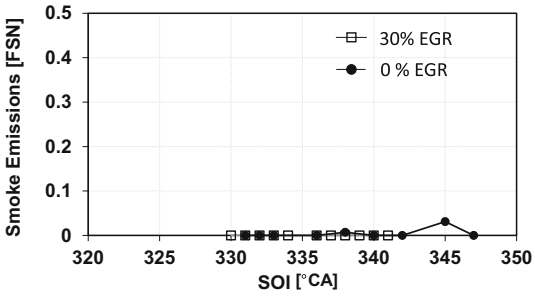


Fig. 23 0% EGR vs. 30% EGR: smoke emissions

Therefore, the high-pressure rise rate of butanol combustion can be mitigated by the application of EGR. A reasonable high efficiency can be maintained in the EGR diluted combustion, though the operation range is narrowed down. The application of EGR is also beneficial for further reducing the NOx emissions.

The principle of the EGR application is that it can dilute the air–fuel mixture and reduce the reactivity of the mixture. From Fig. 15 in the previous section, the air–fuel ratio changes with increased engine load so that the requirement of EGR dilution is also different. As can be seen from Figs. 24 and 25, with a similar amount of EGR and combustion phasing, the increase in engine load results in increased thermal efficiency, while the pressure rise rate also increases. Though there are fluctuations in the trend of efficiency at load level of 6.5 bar as shown in Fig. 24, which is possibly resulted from the slightly higher intake temperature (30 °C compared to 26 °C in other three data

points), the general trend of increased thermal efficiency can be clearly observed. This suggests that a higher load level can tolerate a higher rate of EGR and it requires higher EGR ratio to suppress the higher pressure rise rate. Therefore, it is necessary to carry out an EGR sweep to investigate the effects of increased EGR ratio on the combustion and emission performance.

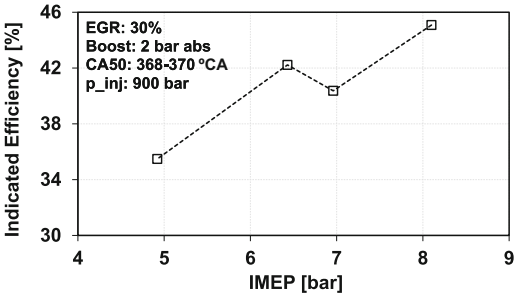


Fig. 24 Indicated efficiency at different engine loads with the same EGR dilution ratio

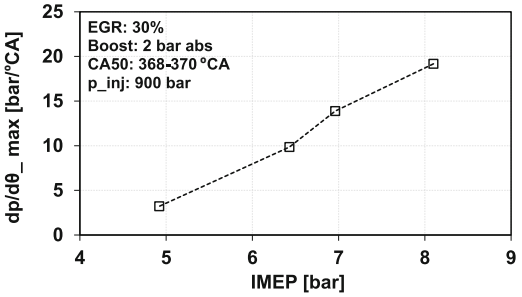


Fig. 25 Peak pressure rise rate at different engine loads with the same EGR dilution ratio

The test conditions are listed in Table 5. In this set of tests, a higher engine load level is used, as higher load has more severe problems of high-pressure rise rate but it can also tolerate a higher EGR ratio. A load level of 10 bar IMEP is realized with a higher injection pressure of 1500 bar. The combustion phasing is kept constant throughout the EGR sweep by adjusting the injection timing.

Table 5 Engine operating conditions for butanol combustion test under different EGR dilution ratios

Nominal IMEP (bar)	10 bar
Injection pressure (bar)	1500
Boost pressure (bar abs)	2
EGR (%)	24–70
Engine speed (rpm)	1500
CA50 (°CA)	365, 375, 380
Injection timing (°CA)	varies

At the engine load of 10 bar IMEP, a minimum EGR ratio of about 20% is required to keep the pressure rise rate within a range of less than 20 bar/°CA for safety concerns. Initially, the combustion phasing is maintained at 365 °CA and the EGR ratio is gradually increased to further reduce the pressure rise rate. However, with the increase of EGR ratio, the smoke emissions are drastically increased as shown in Fig. 26; consequently, EGR ratio higher than 56% is not attempted. As reported in previous sections, butanol generally produces ultralow smoke emissions under lower load levels. The generation of smoke emissions under the higher load condition is due to the change in the air–fuel ratio and in-cylinder temperature conditions. The higher engine load provides the high temperature environment for smoke formation and the increased fuel amount results in less sufficient mixing, both of which increase the chance for smoke formation. When the EGR ratio is low, the formed smoke emissions can still be oxidized; however, when the EGR ratio is further increased, the air is insufficient for the oxidation so that increased smoke emissions are observed. In order to reduce the smoke emissions, the combustion phasing is retarded to 375 °CA and 380 °CA, respectively, to reduce the temperature, thus prolonging the ignition delay for better mixing. With a later combustion phasing, the peak of smoke emissions is lower. The smoke emissions first increase with EGR ratio and gradually decrease when EGR is further increased. The decrease of smoke emission at higher EGR ratio is probably due to the lowered temperature with sufficient dilution, which makes the temperature lower than the threshold temperature of the smoke formation. NOx emissions consistently decrease with increased EGR. Retarding the combustion phasing also benefits the reduction of NOx emissions.

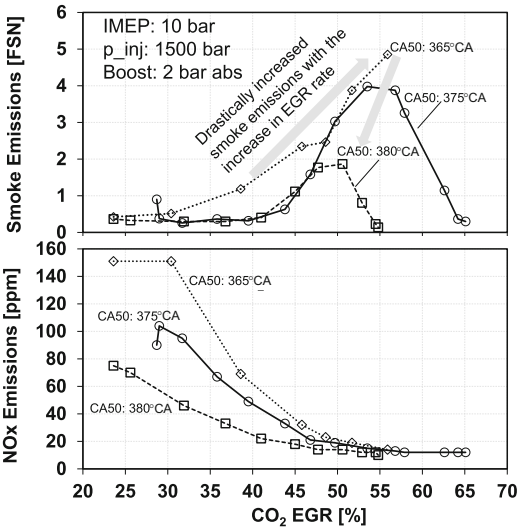


Fig. 26 NOx and smoke emissions under different EGR dilution rates

The pressure rise rate decreases with increased EGR in all the three sets of tests as shown in Fig. 27. The change of combustion phasing does not show obvious impact on the pressure rise rate when the EGR ratio is low. With higher EGR dilution, the later combustion phasing results in much lower pressure rise rate. The indicated efficiency is stable at lower EGR ratios (less than 50%) and drops when the EGR ratio is higher than 50% in all the three sets of tests. The decrease in efficiency is mainly due to the unstable combustion under the high dilution conditions. The effect of combustion phasing on the indicated efficiency is more obvious. With the retardation of the combustion phasing, the indicated efficiency drops significantly from about 42% to 39%.

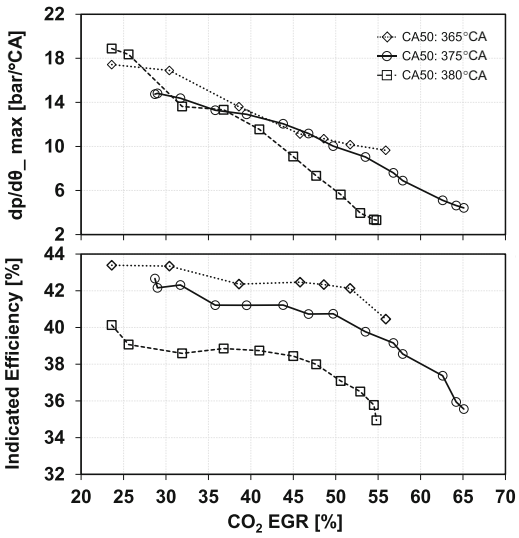


Fig. 27 Peak pressure rise rate and indicated efficiency under different EGR dilution rates

It is concluded from these tests that though EGR is very effective in reducing the peak pressure rise rate and NOx emissions, a higher EGR ratio results in negative side effects of increased smoke emissions and decreased thermal efficiency. A trade-off exists between these requirements. The trade-off will be more significant when the engine load is further increased as the peak pressure rise rate will be even higher and the air–fuel mixture will be even richer, which increases the chance for generating

smoke emissions. The trade-off in pressure rise rate and smoke generation can be alleviated by retarding the combustion phasing, though with the penalty of reduced thermal efficiency.

4 Conclusions

Butanol is a promising alternative fuel to be utilized in internal combustion engines for energy sustainability and greenhouse gas reduction. The butanol combustion also exhibits low smoke and NOx emissions. The combustion characteristics of butanol are significantly different from those of conventional diesel fuel due to the changes in fuel properties. Engine tests are carried out under different load levels to investigate the trend of butanol combustion with elevated engine load. EGR is applied to assist the control of butanol combustion. Following conclusions can be drawn from the test results:

1. Butanol combustion with direct injection has a much narrower window of fuel injection timing for stable operation compared to diesel combustion. The constraints are mainly coming from two aspects: the high-pressure rise rate when the rapid burning of a premixed butanol–air mixture occurs close to TDC, and the unstable combustion when the late combustion happens in the expansion stroke resulted by either long ignition delay or late injection timing.
2. The ignition delay of butanol is much longer than that of diesel fuel, and it reduces when the injection is gradually retarded toward TDC. The combustion phasing of butanol is not linearly related to the injection timing. A very late combustion phasing may be resulted when the injection timing is scheduled too early, due to the longer ignition delay time.
3. Butanol combustion produces significantly lower NOx and smoke emissions, though the THC and CO emissions are generally higher.
4. Constrained by the excessive pressure rise rate, the operation window of butanol combustion is narrower when the engine load is increased from 2 bar to 7.5 bar under zero percent EGR. Near-zero smoke emissions are produced in all the cases, while NOx emissions slightly increase with engine load.
5. With the application of EGR, the pressure rise rate and NOx emissions are effectively reduced.
6. Though EGR is effective in reducing the pressure rise rate, increased EGR ratio results in increased smoke emissions under higher engine loads. The trade-off between pressure rise rate and smoke emissions limited the further increase in engine load with the single-shot injection strategy.

Acknowledgments. This research is supported by NSERC CRD, APC, Discovery, CRC, CREATE programs; the CFI-ORF New Initiative Program; the NCE BioFuelNet programs; the Ford Motor Company, and the University of Windsor.

References

1. U.S Energy Information Administration: Annual Energy Outlook (2018). www.eia.gov/eao. Accessed 02 May 2018
2. Salvi, B.L., Subramanian, K.A., Panwar, N.L.: Alternative fuels for transportation vehicles: a technical review. Renew. Sustain. Energy Rev. **25**, 404–419 (2013)
3. Alternative Fuels Data Center: Maps and Data. http://www.afdc.energy.gov/data/. Accessed 02 May 2018
4. Han, X., Zheng, M., Wang, J.: Fuel suitability for low temperature combustion in compression ignition engines. Fuel **109**, 336–349 (2013)
5. Lu, X., Han, D., Huang, Z.: Fuel design and management for the control of advanced compression-ignition combustion modes. Prog. Energy Combust. Sci. **37**(6), 741–783 (2011)
6. Dec, J.E.: Advanced compression-ignition engines—understanding the in-cylinder processes. Proc. Combust. Inst. **32**(2), 2727–2742 (2009)
7. Rakopoulos, D.C., Rakopoulos, C.D., Papagiannakis, R.G., Kyritsis, D.C.: Combustion heat release analysis of ethanol or n-butanol diesel fuel blends in heavy-duty DI diesel engine. Fuel **90**(5), 1855–1867 (2011)
8. Rakopoulos, D.C., Rakopoulos, C.D., Giakoumis, E.G., Dimaratos, A.M., Kyritsis, D.C.: Effects of butanol–diesel fuel blends on the performance and emissions of a high-speed DI diesel engine. Energy Convers. Manag. **51**(10), 1989–1997 (2010)
9. Rakopoulos, D.C., Rakopoulos, C.D., Hountalas, D.T., Kakaras, E.C., Giakoumis, E.G., Papagiannakis, R.G.: Investigation of the performance and emissions of bus engine operating on butanol/diesel fuel blends. Fuel **89**(10), 2781–2790 (2010)
10. Choi, B., Jiang, X., Kim, Y.K., Jung, G., Lee, C., Choi, I., Song, C.S.: Effect of diesel fuel blend with n-butanol on the emission of a turbocharged common rail direct injection diesel engine. Appl. Energy **146**, 20–28 (2015)
11. Abu-Qudais, M., Haddad, O., Qudaisat, M.: The effect of alcohol fumigation on diesel engine performance and emissions. Energy Convers. Manag. **41**(4), 389–399 (2000)
12. Rajesh Kumar, B., Saravanan, S.: Effects of iso-butanol/diesel and n-pentanol/diesel blends on performance and emissions of a DI diesel engine under premixed LTC (low temperature combustion) mode. Fuel **170**, 49–59 (2016)
13. Hansen, A.: Ethanol-diesel fuel blends - a review. Bioresour. Technol. **96**(3), 277–285 (2005)
14. Xing-cai, L., Jian-guang, Y., Wu-gao, Z., Zhen, H.: Effect of cetane number improver on heat release rate and emissions of high speed diesel engine fueled with ethanol–diesel blend fuel. Fuel **83**(14–15), 2013–2020 (2004)
15. Iannuzzi, S.E., Valentino, G.: Comparative behavior of gasoline–diesel/butanol–diesel blends and injection strategy management on performance and emissions of a light duty diesel engine. Energy **71**, 321–331 (2014)
16. Chen, Z., Liu, J., Han, Z., Du, B., Liu, Y., Lee, C.: Study on performance and emissions of a passenger-car diesel engine fueled with butanol–diesel blends. Energy **55**, 638–646 (2013)

17. Maurya, R.K., Agarwal, A.K.: Experimental investigation on the effect of intake air temperature and air–fuel ratio on cycle-to-cycle variations of HCCI combustion and performance parameters. Appl. Energy **88**(4), 1153–1163 (2011)
18. Saxena, S., Vuilleumier, D., Kozarac, D., Krieck, M., Dibble, R., Aceves, S.: Optimal operating conditions for wet ethanol in a HCCI engine using exhaust gas heat recovery. Appl. Energy **116**, 269–277 (2014)
19. Zheng, M., Li, T., Han, X.: Direct injection of neat n-butanol for enabling clean low temperature combustion in a modern diesel engine. Fuel **142**, 28–37 (2015)
20. Bahri, B., Aziz, A.A., Shahbakhti, M., Muhamad Said, M.F.: Understanding and detecting misfire in an HCCI engine fuelled with ethanol. Appl. Energy **108**, 24–33 (2013)
21. Sjöberg, M., Dec, J.E.: Effects of EGR and its constituents on HCCI auto ignition of ethanol. Proc. Combust. Inst. **33**(2), 3031–3038 (2011)
22. Zheng, M., Han, X., Asad, U., Wang, J.: Investigation of butanol-fuelled HCCI combustion on a high efficiency diesel engine. Energy Convers. Manag. **98**, 215–224 (2015)
23. Maurya, R.K., Agarwal, A.K.: Experimental investigations of performance, combustion and emission characteristics of ethanol and methanol fueled HCCI engine. Fuel Process. Technol. **126**, 30–48 (2014)
24. Maurya, R.K., Agarwal, A.K.: Experimental study of combustion and emission characteristics of ethanol fuelled port injected homogeneous charge compression ignition (HCCI) combustion engine. Appl. Energy **88**(4), 1169–1180 (2011)
25. Liu, M.-B., He, B.-Q., Zhao, H.: Effect of air dilution and effective compression ratio on the combustion characteristics of a HCCI (homogeneous charge compression ignition) engine fuelled with n-butanol. Energy **85**, 296–303 (2015)
26. Jeftić, M., Yang, Z., Reader, G.T., Zheng, M.: Fuel efficiency analysis and peak pressure rise rate improvement for neat n-butanol injection strategies. Proc. Inst. Mech. Eng. Part D: J. Automob. Eng. **231**(1), 50–65 (2017)
27. Han, X., Yang, Z., Wang, M., et al.: Clean combustion of n-butanol as a next generation biofuel for diesel engines. Appl. Energy **198**, 347–359 (2017)
28. Yu, S., Zheng, M.: Ethanol–diesel premixed charge compression ignition to achieve clean combustion under high loads. Proc. Inst. Mech. Eng. D J. Automob. Eng. **230**(4), 527–541 (2015)
29. Zheng, C., Liu, J., Wu, Z., Lee, C.: Effects of port fuel injection (PFI) of n-butanol and EGR on combustion and emissions of a direct injection diesel engine. Energy Convers. Manage. **76**, 725–731 (2013)
30. Zhu, Y., Zheng, C., Liu, J.: Emission, efficiency, and influence in a diesel n-butanol dual-injection engine. Energy Convers. Manage. **87**, 385–391 (2014)

Nozzle Entry Effects and Cavitation Inception in Crossflow Hydroturbines

R. C. Adhikari and D. H. Wood(✉)

Department of Mechanical and Manufacturing Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada
dhwood@ucalgary.ca

Abstract. Crossflow hydroturbines are simple and cheap to manufacture. This work is part of a program to improve the design methodology of small crossflow hydroturbines for remote power systems in developing countries. Adhikari and Wood (2018) showed that the exit arc—the circumferential extent of the flow exiting the runner—is typically half the length of the entry arc, θ_s, of the flow entering the runner. Negative gauge pressures are thus required in the exiting flow to achieve high efficiency. Here, we investigate increasing θ_s to reduce the negative pressures and the influence on cavitation which begins on the blades near the exit. The influence of θ_s is studied using computational fluid dynamics simulation of two turbines whose performance has been measured experimentally. No evidence of cavitation was found in the 0.53 kW turbine (efficiency η = 88%, flow rate Q = 46 lps, and head H = 1.337 m) at any operating speed, whereas significant cavitation was found at maximum efficiency (η = 91%) in the improved design of the 7 kW turbine (original η = 69%, Q = 105 lps, and H = 10 m). The extent of cavitation decreased as θ_s was increased but this also decreased η from 91% to 87%. Changes in radius ratio and inner blade angle decreased significantly the size of cavitation inception region with only a small decrease in maximum efficiency. Finally, redesigning the turbine for reduced H = 8 m and Q = 94 lps avoided cavitation while achieving a similar maximum efficiency of 89.3%. Nevertheless, it may be possible to avoid cavitation through investigation of an optimum combination of geometrical parameters but this would require a considerable computational effort which probably should be guided by experiment.

Keywords: Crossflow turbine · Efficiency · Entry arc · RANS simulation Cavitation inception

Nomenclature

β_1	Runner entry flow angle (rad or degrees)
β_{1b}	Outer blade angle (rad or degrees)
β_{2b}	Inner blade angle (rad or degrees)
H	Turbine head (m)
h_0	Nozzle throat (m)

A. Vasel and D. S.-K. Ting (Eds.): EAS 2018, SPE, pp. 80–92, 2019.
https://doi.org/10.1007/978-3-030-00105-6_5

$h(\theta)$ Nozzle rear wall from the runner tip, $R(\theta)$–R_1, (m)
N_b Number of blades
Q Flow rate through turbine (lps)
Q_{max} Maximum (design) flow rate (lps)
$R(\theta)$ Radius of nozzle rear wall (m)
R_1 Outer radius of the runner (m)
R_2 Inner radius of the runner (m)
p Local static pressure (Pa)
p_v Saturated vapor pressure (Pa)
U_0 Nozzle inlet velocity (m/s)
W Nozzle and runner width (m)
$\dot{W}$ Turbine power (kW)
ω Runner angular speed (m/s)
θ_s Entry arc (rad or degrees)
η Turbine efficiency (%)
ρ Water density (kg/m^3)

1 Introduction

Many small, remote hydropower systems in developing countries employ crossflow turbines due to their inherent simplicity of manufacture and low cost. Other turbine types such as Pelton and Francis, however, can easily achieve a maximum efficiency above 90% operating at moderate to high head H, whereas crossflow turbines typically reach less than 85% at low to moderate head. The work described here is part of a program to improve the efficiency of crossflow turbines while maintaining their simplicity and low cost.

As can be inferred from Fig. 1, the turbine nozzle accelerates the flow. Water then enters the runner at a flow angle β_1 that must match the blade inlet angle β_{1b}. Adhikari and Wood (2017) showed that the nozzle should be designed to convert H into kinetic energy at the runner entry. This principle leads to a simple equation for the nozzle geometry, but does not fix the nozzle entry arc θ_s. Adhikari (2016) used this methodology to redesign the 7 kW turbine of Dakers and Martin (1982) with a measured efficiency $\eta = 69\%$ to achieve $\eta = 91\%$. The original turbine had $\theta_s = 69^\circ$ which was altered to 80° during the redesign. The 0.53 kW turbine of Desai (1993) with a maximum $\eta = 88\%$ could only be improved marginally, largely because the original nozzle effectively converted H to kinetic energy. Desai's turbine had $\theta_s = 90^\circ$. Neither reference explained the choice of θ_s, and no systematic study of the effect of θ_s on performance is available in the open literature. We show in the next section that the height of the nozzle throat, h_0, must be altered if θ_s is changed.

The crossflow turbine is unique in that the flow passes twice through the runner. Only a part of the power is extracted by the "first stage" in Fig. 1. Then the water traverses the inner airspace and enters the "second stage" where the remainder of the power is extracted. Approximately, 25% of the blades interact with the water at any time, so the remaining section of the runner is exposed to ambient air. It is necessary to undertake detailed computational fluid dynamics simulations of the turbine to determine the relative power extraction of the two stages as it would be almost impossible to do this experimentally. If the turbine is efficient, then no angular momentum leaves the second stage (Adhikari and Wood 2018), which, in turn, constrains the exit arc; $\theta_e \approx 0.5\theta_s$. This has at least two consequences. First, if $\theta_s \approx 90^\circ$, then the angular extent of the entering and exiting flow is less than 180° and the simulations of Adhikari and Wood (2018) show that the two stages are contiguous. This allows the possibility of building dual nozzle crossflow turbines to double the power density. Second, the exit radial velocity is high and so the only way that a small amount of energy leaves the runner is for the gauge pressure in the second stage to be negative. This raises the possibility of cavitation at high efficiency. Cavitation in crossflow turbines has only been studied by Adhikari et al. (2016) in their simulations of the inefficient 7 kW turbine. As expected, cavitation started on the suction surface of the blades near the entry to the second stage.

The aim of this paper is to explore the effects of varying θ_s on crossflow turbine performance and to investigate cavitation at conditions around maximum efficiency. We used steady Reynolds-averaged Navier–Stokes (RANS) simulations with two-phase homogeneous flow with free-surface effects between water and air, in which RANS equations were solved for both fluids. The SST $k - \omega$ turbulence model provided satisfactory results for predicting the turbine performance, which was validated through comparison with experimental results. The details of the CFD and experimental results can be found in Adhikari and Wood (2017). The next section describes the changes required to accommodate a change in θ_s and the method by which the inception of cavitation is determined. The following section describes the effects of changing θ_s on the performance of the 0.53 kW turbine for which no cavitation was found, and the improved 7 kW turbine in which cavitation occurs in the second stage at maximum efficiency. The final section contains the summary and conclusions.

2 Nozzle Design and Size of the Entry Arc

The simple equations of Adhikari and Wood (2017) assume uniform flow in the nozzle of width W, which is the same width as the nozzle and runner, and uniform velocity at the throat, U_0. With an outer blade radius, R_1, the conversion of H to kinetic energy at runner entry requires

$$gH = \frac{U_0^2}{2}\left(1 + \frac{h_0^2}{R_1^2\theta_s^2}\right) \tag{1}$$

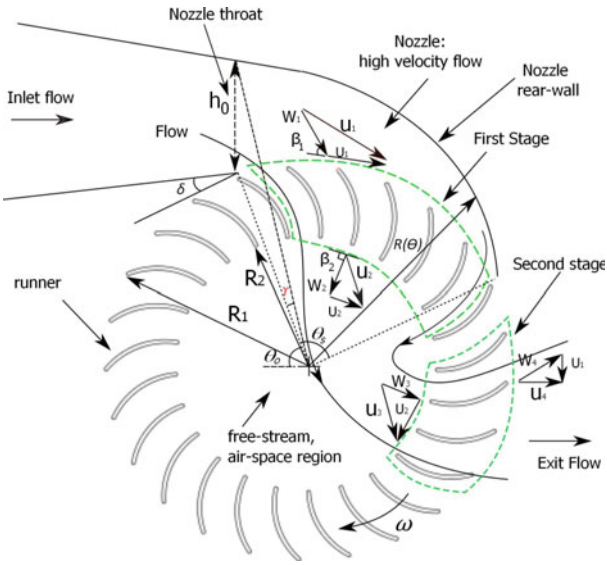


Fig. 1. Schematic illustration of key geometric parameters of crossflow turbine

which can be rewritten in terms of the flow rate Q, as

$$\frac{2WgH}{Q^2} = \frac{1}{h_0^2} + \frac{1}{R_1^2\theta_s^2} \tag{2}$$

All terms on the left are independent of θ_s, so h_0 must be altered if θ_s changes. This, and the associated change in the nozzle rear wall shape, $h(\theta)$, was done in a series of simulations of the 0.53 kW turbine for $\theta_s = 80^\circ$, 85°, 90° and 105°.

The influence of θ_s on the performance of the 0.53 kW turbine was investigated by examining the runner total entry velocity, flow angle, and the percentage power extraction in the first and second stages. The onset of cavitation was also investigated for both turbines. For the 7 kW turbine of Dakers and Martin (1982), the number of blades, N_b, and the outer and inner blade angles were altered as described in Adhikari and Wood (2018) to increase η from 69% to 91%. The computational study of cavitation by Adhikari et al. (2016) considered only the original 7 kW turbine, which was experimentally studied by Dakers and Martin (1982), and we are aware of no other investigations of cavitation for crossflow turbines. In Adhikari et al. (2016) and here, cavitation occurs when the local pressure, p, falls to the saturated vapor pressure of water, p_v. A further reduction of p increases cavitation but the development of cavitation requires a more sophisticated multiphase flow model than the one used here.

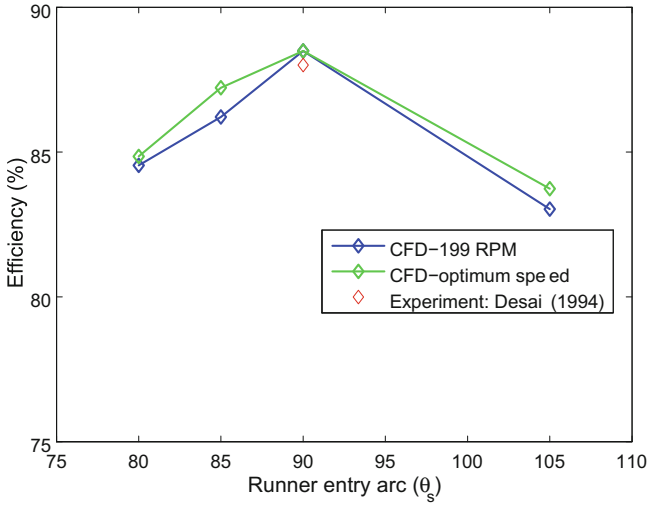


Fig. 2. Influence of runner entry arc angle θ_s on runner efficiency. The change in θ_s for the same runner changes the runner entry velocity, flow angle, and the number of blades extracting the power in the first and second stages

3 Results and Discussion

3.1 0.53 KW Turbine

Computations were conducted for θ_s = 80, 85, 90, and 105°. The original turbine had θ_s = 90°, at which η = 88% (Desai 1993). h_0 was adjusted according to Eq. (2). This change alters the relationship between the radial and circumferential velocities at runner entry which should be matched by a change in the blade inlet angle. No change, however, was made to the geometry of the runner for these initial calculations. Alternatively, ω_{opt} for maximum efficiency should be changed; from Eq. (2) of Adhikari and Wood (2018), $\omega_{opt} \sim h_0^2$. Calculations were conducted at the experimentally and computationally determined ω_{opt} = 199.1 RPM for θ_s = 90°, which are shown as the blue diamonds in Fig. 2. The green diamonds show η with the speed altered by Eq. (2). For each θ_s, the rear wall profile $R(\theta)$ was determined using the nozzle design methodology presented in Adhikari and Wood (2017). It is noted that the comparison between the CFD and experimental measurements could be made only for the original θ_s = 90° (Desai 1993). It is evident that the effect of θ_s on runner efficiency is significant over the range θ_s = 80-105° and that the designers of the 0.53 kW turbine made an excellent choice for θ_s and the matching runner geometry. Unfortunately, Desai (1993) gave no reasons for that decision. The efficiency can be examined in terms of total inlet velocity, flow angle β_1, and percentage stage power as follows.

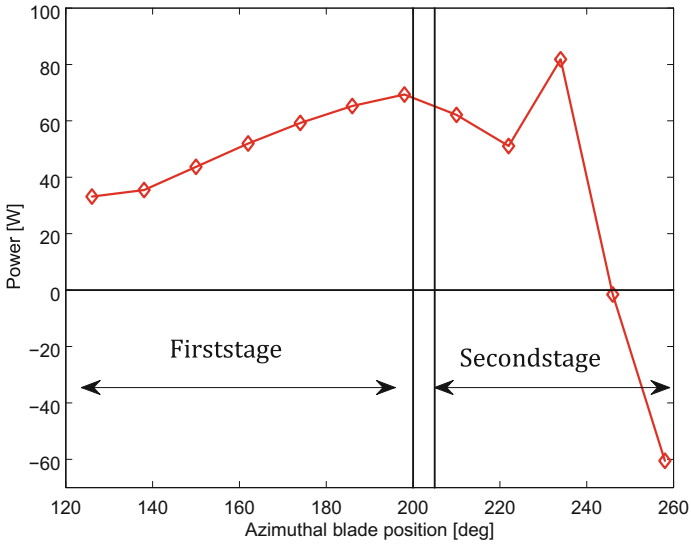


Fig. 3. Stage performance at entry arc angle $\theta_s = 80^\circ$. There is a significant portion of negative power at the second stage

The power extracted by the blades in the two stages is shown in Fig. 3 for $\theta_s = 80^\circ$. Significant negative power was found in the second stage, and the first-stage extraction is lower compared to the case at $\theta_s = 90^\circ$ shown in Fig. 8 of Adhikari and Wood (2018). This is the main reason for the lower performance of the runner.

Figure 4 compares the percentage power extraction in the first and second stages for different θ_s. At the maximum η, for $\theta_s = 90^\circ$, the first stage extracted about 69% of total power. For $\theta_s = 105^\circ$, the first-stage power extraction has increased but the maximum runner efficiency has decreased. Similarly, for $\theta_s = 120^\circ$, the first-stage power as well as the maximum runner efficiency has decreased significantly. For $\theta_s < 90^\circ$, the first-stage power has increased as θ_s decreases. This implies that there is an optimum θ_s for the first-stage power extraction at which the runner achieves maximum efficiency. These results demonstrate that a particular value of θ_s maximizes the runner performance but it must be optimized in conjunction with other design parameters, principally the inlet blade angle, to achieve maximum runner efficiency. In other words, there is a particular θ_s at which the total inlet velocity is approximately constant at the value representing the total conversion of H into kinetic energy, and where β_1 matches closely to the outer blade angle β_{1b}.

Figures 5 and 6 show the water velocity contours for $\theta_s = 80^\circ$ and 105°, respectively. There is no flow separation for the lower θ_s but there is a significant separation in the first stage for the larger θ_s. The reason can be seen from Figs. 7 and 8 which compare β_1 to β_{1b}. For $\theta_s = 105^\circ$, β_1 is too large for most of the entry arc. Interestingly, there is little difference in θ_e for the large change in θ_s between these two figures. It is likely that the reduced η (relative to $\theta_s = 80^\circ$) is the reason.

To determine whether the flow in the runner is cavitating or not, the absolute pressure was plotted at maximum efficiency points. In all the cases considered for the

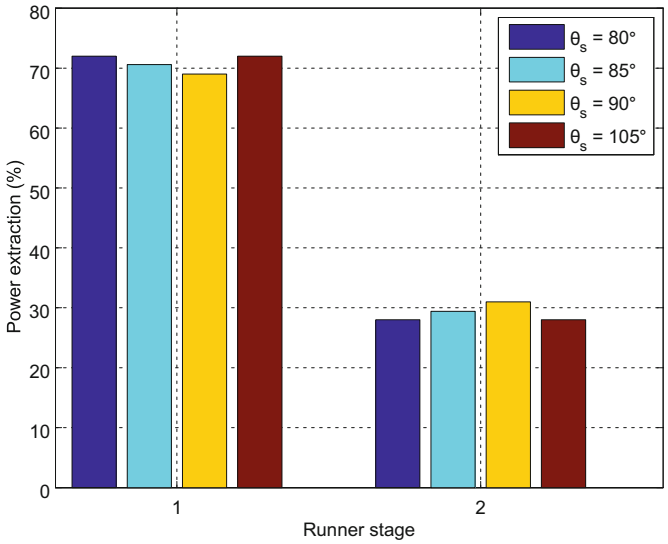


Fig. 4. Influence of runner entry arc angle θ_s on runner stage performance at maximum efficiency. Maximum efficiency occurs at $\theta_s = 90^\circ$

0.53 kW turbine, the lowest absolute pressure in the runner was above 58 kPa, whereas cavitation occurs at the water vapor pressure of 3.17 kPa at 25 °C. So the runner is free of cavitation in all the cases. It is, however, interesting to note that as θ_s increased from 90°, the lowest absolute pressure in the runner decreased slightly.

3.2 Improved 7 KW Turbine—Cavitation

The 7 kW turbine with a maximum efficiency of 69%, measured by Dakers and Martin (1982), was improved by Adhikari (2016) and Adhikari and Wood (2018) using numerical simulations with reference to the good practice of the 0.53 kW turbine investigated above. The improved version of the 7 kW turbine has a maximum efficiency of 91%. Adhikari et al. (2016) showed that cavitation in the original turbine occurred near the entrance to the second stage; the details of which are given in Adhikari et al. (2016). Cavitation in the improved turbine is discussed here. Figure 9 shows the important finding that cavitation begins near the leading edge of three blades in the second stage at maximum efficiency ($\eta = 91\%$), which is similar to that of the original turbine. The mean water velocity vectors, not shown here, indicated that cavitation started within a region of separated flow. In practice, it is essential to avoid cavitation and maintain a high maximum efficiency. We investigated the influence of a number of turbine geometrical parameters on cavitation inception, while attempting to avoid a reduction in the maximum efficiency (91%). The parameters investigated are the entry arc θ_s, number of blades N_b, radius ratio R_2/R_1, inner blade angle β_{2b}, and outer radius R_1. The results are briefly summarized here.

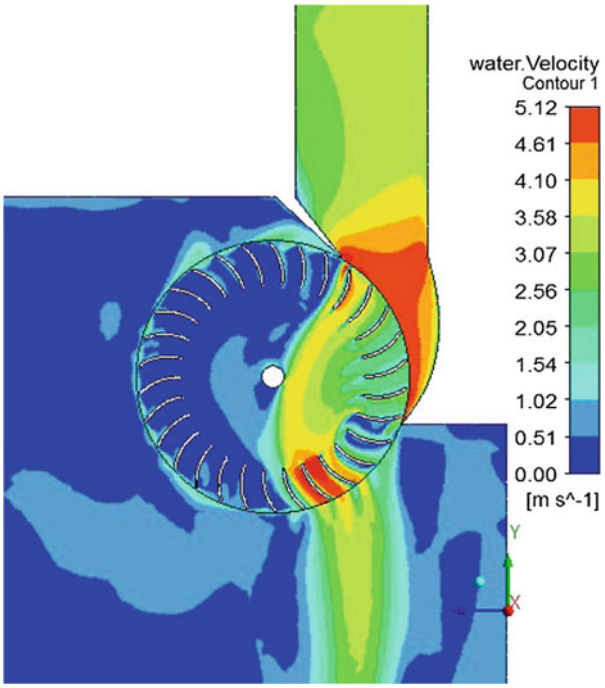


Fig. 5. Contours of mean water velocity at $\theta_s = 80^\circ$. There is no flow separation on the runner blades

Since θ_s influences cavitation as discussed above for the 0.53 kW turbine, θ_s was increased from the optimum value of 80° to 90°. $\theta_s = 90^\circ$ decreased the size of incipient cavitation region as shown in Fig. 10, but this was accompanied by a decrease in the maximum efficiency, from 91% to 89.7%.

As N_b influences separation from the blades, and by inference, the inception of cavitation, N_b, was reduced from 30 to 20. As expected, cavitation was considerably reduced by being confined to one blade, but not fully avoided as shown in Fig. 11. Also, the efficiency decreased to 87%.

Similarly, as the inner to outer radius ratio R_2/R_1 is related to blade chord length and radius, it should have an influence on cavitation inception. So R_2/R_1 was increased from the original optimum value of 0.67 to 0.68, which significantly decreased the size of incipient cavitation region, in comparison to θ_s or N_b, but did not fully avoid cavitation. In the interests of brevity, these results are not shown. More importantly, the efficiency decreased only marginally from 91% to 90.3%.

In the next series of simulations, for $R_2/R_1 = 0.68$, β_{2b} was increased from 90 to 92.5°. The objective was to minimize the flow separation at the entry to the second stage by reducing the difference between the flow angle and β_{2b}. This combination of

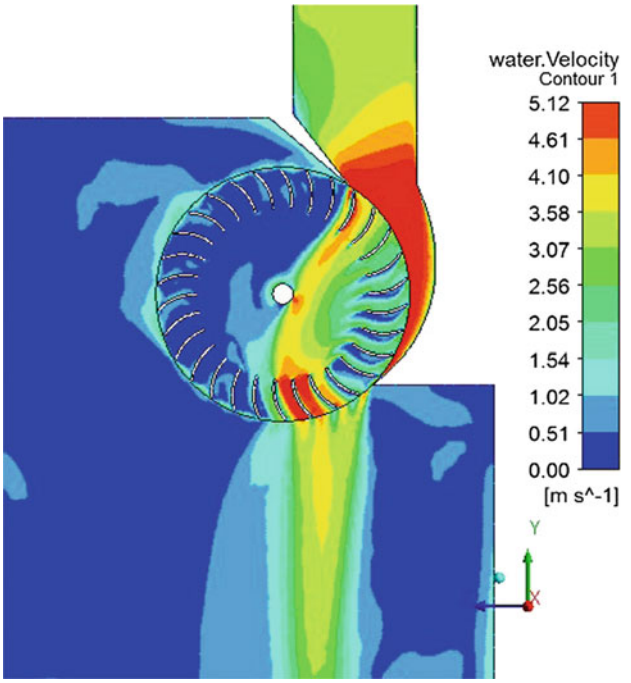


Fig. 6. Contours of mean water velocity at $\theta_s = 105^\circ$. There is a significant flow separation on the suction sides of the first-stage blades

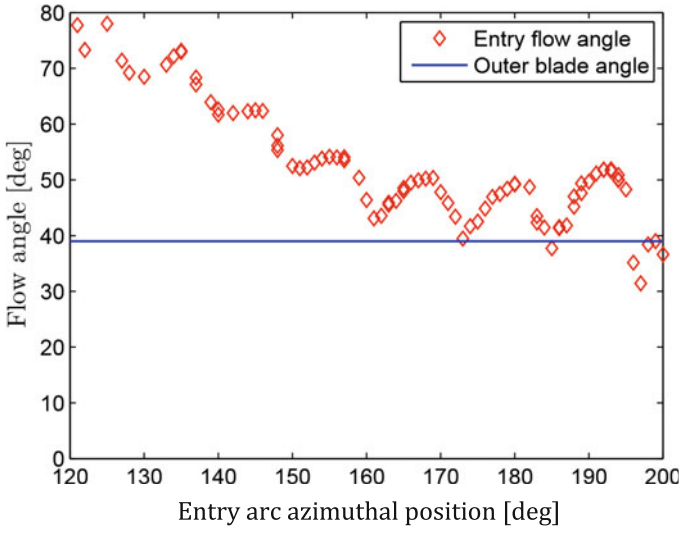


Fig. 7. Entry flow angle β_1 compared to the outer blade angle β_{1b} at $\theta_s = 80^\circ$

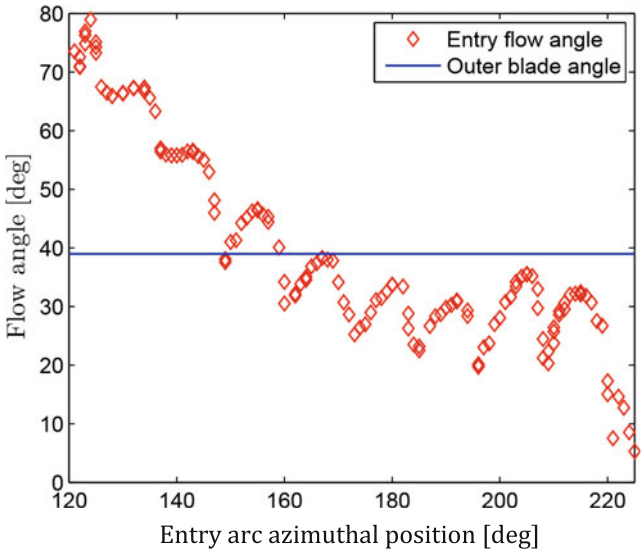


Fig. 8. Entry flow angle β_1 compared to the outer blade angle β_{1b} at $\theta_s = 105^\circ$

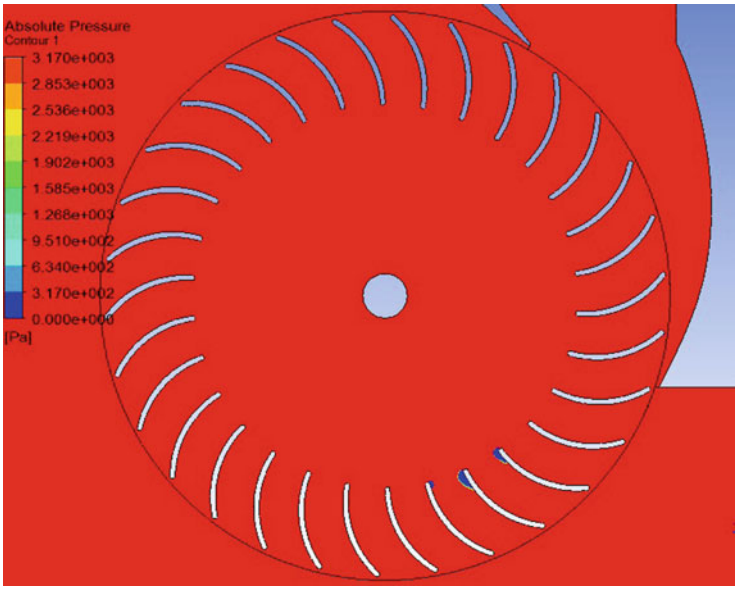


Fig. 9. Cavitation at maximum efficiency [$\eta = 91\%$, $\omega_{opt} = 500$ RPM, $\theta_s = 80^\circ$]. The pressure contours up to the water vapor pressure 3.17 kPa show the extent of incipient cavitation near the entry of second stage blades

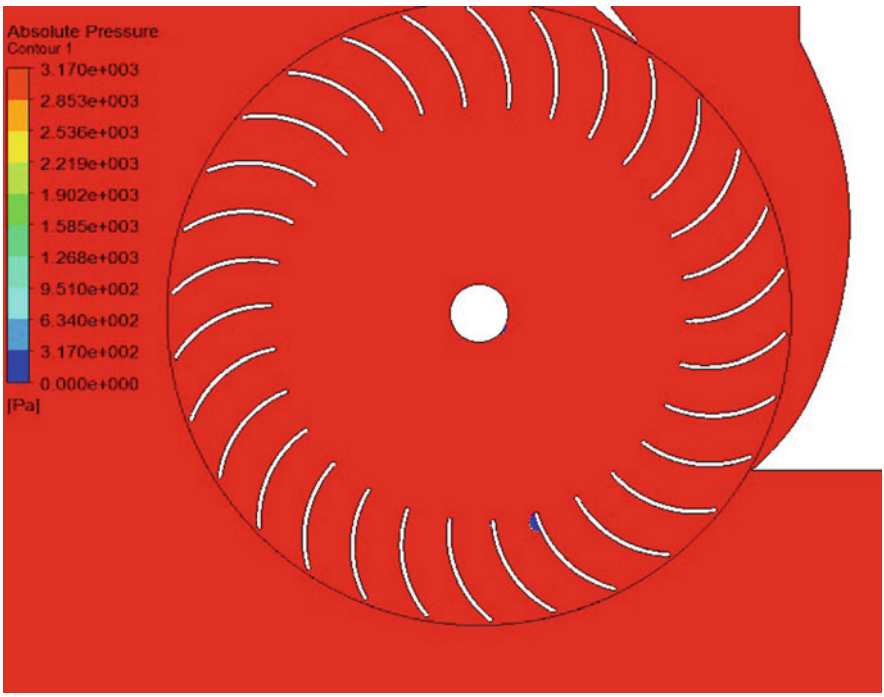


Fig. 10. Cavitation at maximum efficiency [N = 500 RPM, θ_s = 90°]. The pressure contours up to the water vapor pressure 3.17 kPa show the extent of incipient cavitation near the entry of second stage blades

changes to the design parameters confined cavitation to a smaller region than any of the above actions, with only a small decrease in efficiency, from 91 to 90.1%. However, cavitation was not fully eliminated. Likewise, when R_1 was increased from 158 to 160 mm, the size of cavitation region was not altered.

The above results indicate that totally avoiding cavitation while maintaining maximum efficiency could not be achieved. It may be possible to avoid cavitation through investigation of a combination of changes to the geometrical parameters but this would require a considerable computational effort which probably should be guided by experiment. Alternatively, Adhikari et al. (2016) showed that reducing the head H from 10 to 8 m avoided cavitation inception while maintaining the similar maximum efficiency in the original 7 kW turbine. Therefore, simulations were conducted by modifying the turbine design for H = 8 m and Q = 94 lps, following the nozzle design method outlined in Sect. 2, without altering the runner design. Cavitation did not occur at maximum efficiency, which was 89.3%, only slightly lower than the baseline design (91%).

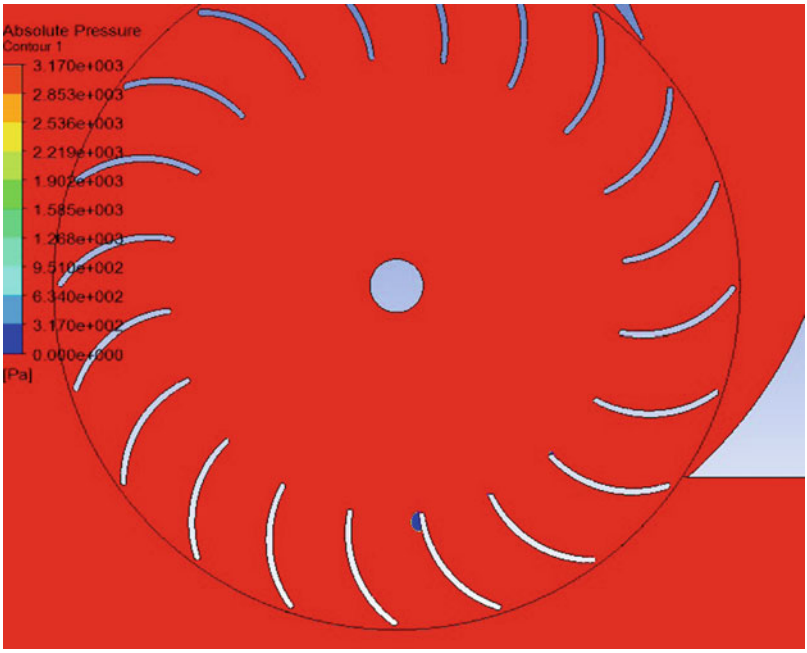


Fig. 11. Improved 7 kW turbine: cavitation on a blade at the second stage for $N_b = 20$ (maximum efficiency = 87%)

4 Summary and Conclusions

This chapter shows that there is a complex interaction between some of the important design parameters of crossflow hydroturbines, the efficiency and tendency toward cavitation. Two turbines were studied: a 0.53 kW one with a measured maximum efficiency of 88% and a 7 kW whose original efficiency was 69%. By a combination of computational fluid dynamics analysis and reference to the more efficient turbine, we were able in previous work to redesign the nozzle and runner to increase the efficiency of 7 kW turbine to 91%. The 0.53 kW turbine always operated well away from cavitation and this was not changed when the entry arc of the nozzle was altered with the aim of altering the negative gauge pressures in the second stage. Without adjusting the runner geometry or the operational speed, the efficiency of the 0.53 kW turbine reduced when the entry arc was altered from the original value of 90°.

We have shown for the first time that cavitation can occur in a turbine operating at a maximum efficiency of around 91%. The interaction of the number of blades, entry arc, radius ratio, and inner blade angle with the extent of cavitation inception and maximum efficiency was studied. It is shown that reducing the number of blades reduces the extent of cavitation but also lowers the maximum efficiency. Changes in radius ratio and inner blade angle decreased significantly the size of cavitation inception region with only a small decrease in maximum efficiency. Finally, redesigning the turbine for reduced head and flow rate avoided cavitation while achieving a similar maximum efficiency. Nevertheless, it is emphasized here that it may be possible to avoid

cavitation through a detailed investigation of the optimum combination of geometrical parameters, but it requires a considerable amount of computational effort. It is likely that experimental studies are also needed.

Acknowledgement. The authors acknowledge the funding from the Schulich Research Chair in Renewable Energy at the University of Calgary, Canada. We also acknowledge WestGrid Canada for providing high-performance computers to perform flow simulations.

References

Adhikari, R.: Design improvement of crossflow hydro turbine. Ph.D. thesis University of Calgary, Canada (2016)

Adhikari, R., Wood, D.: A new nozzle design methodology for high efficiency crossflow hydro turbines. Energy. Sustain. Dev. **41**, 139–148 (2017)

Adhikari, R., Wood, D.: Computational analysis of a double-nozzle crossflow turbines. Energies, (submitted) (2018)

Adhikari, R.C., Vaz, J., Wood, D.: Cavitation inception in crossflow hydro turbines. Energies **9**, 237–248 (2016)

Adhikari, R.C., Wood, D.: Design of high efficiency crossflow turbines: review and extension. Energies **11**, 267 (2018). https://doi.org/10.3390/en11020267

Dakers, A., Martin, G.: Development of a simple cross-flow water turbine for rural use. In: Agricultural Engineering Conference 1982: Resources, Efficient Use and Conservation; Preprints of Papers, p. 35. Institution of Engineers, Australia (1982)

Desai, V.R.: A Parametric study of the cross-flow turbine performance. Ph.D. thesis Clemson University, USA (1993)

Compact CPV—Sustainable Approach for Efficient Solar Energy Capture with Hybrid Concentrated Photovoltaic Thermal (CPVT) System and Hydrogen Production

Muhammad Burhan(✉), Muhammad Wakil Shahzad, and Kim Choon Ng

Water Desalination and Reuse Centre (WDRC), Biological and Environmental Science & Engineering (BESE), King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
muhammad.burhan@kaust.edu.sa

Abstract. Solar energy being intermittent in nature can provide a sustainable, steady, and high-density energy source when converted into electrolytic hydrogen. However, in the current photovoltaic market trend with 99% conventional single junction PV panels, this cannot be achieved efficiently and economically. The advent of the multi-junction solar cells (MJCs), with cell efficiency exceeding 46%, has yet to receive widespread acceptance in the current PV market in form of concentrated photovoltaic (CPV) system, because of its system design complexity, limiting its application scope and customers. The objective of this paper is to develop a low-cost compact CPV system that will not only eliminate its application and installation-related restrictions but it is also introducing a highly efficient and sustainable photovoltaic system for common consumer, to convert intermittent sunlight into green hydrogen. The developed CPV system negates the common conviction by showing two times more power output than the flat plate PV, in the tropical region. In addition, sunlight to hydrogen conversion efficiency of 18% is recorded for CPV, which is two times higher than alone electricity production efficiency of flat plate PV. As concentrated photovoltaic (CPV) system can operate at $\times 1000$ concentration ratio, therefore, such high concentration ratio requires heat dissipation from the cell area to maintain optimum temperature. With such heat recovery, the hybrid CPVT system has shown solar energy conversion efficiency of 71%.

1 Introduction

The global warming circumstance of 1.5 °C rise in the average ambient temperature [1–5] is deemed inevitable due to the excessive emission of greenhouse gases such as carbon dioxide [2]. Today, the renewable energy contributes only about 6.7% of the global energy demand [3]. The reason for the slow implementation of renewable energy sources can be attributed to their vulnerability to intermittency, low-energy density, and localized availability that makes these energy sources mediocre as compared to the conventional fossil fuel-based sources. Among all renewable sources,

A. Vasel and D. S.-K. Ting (Eds.): EAS 2018, SPE, pp. 93–102, 2019.
https://doi.org/10.1007/978-3-030-00105-6_6

the solar energy has the highest potential [4] in meeting a significant portion of the total energy demand of future and for such energy sources to achieve sustainability, its energy quality and supply continuity are to be improved [5].

Current photovoltaic (PV) systems for electricity generation, available hitherto, comprise the stationary single junction silicon-based solar cells such as monocrystalline, polycrystalline, and thin films [6, 7] with more than 98% share of photovoltaic market. Although such cells have a theoretical limit of 31% efficiency, the actual long-term electricity rating is merely less than 8% [8]. The advent of the third-generation multi-junction solar cells (MJCs), with the instantaneous cell efficiency exceeding 46% [9] (but the quantum limit is 86%), has yet to receive widespread acceptance in form of concentrated photovoltaic (CPV) system, in the current PV market because of its system design complexity in application. Conventional gigantic (big size) CPV systems require massive structures and open land spaces, limiting its application scope and customers. It is not simple to have compact CPV design as it demands more units of solar trackers or control hardware and they demand very high degree tracking accuracy as high as 0.1^{o}. Therefore, if the cost of the solar tracking units cannot be reduced, then the compact CPV concept cannot compete with the conventional gigantic (big size) CPV design, however, without compromising on the required tracking accuracy.

On the other hand, despite high-grade energy production as electricity, it is still intermittent and requires energy storage for steady power supply, while electrochemical storage in form of battery does not make solar energy as a viable option to replace fossil fuels [10]. Electrolytic production of hydrogen by splitting water using solar electricity not only provides a portable green energy but also acts as high-density energy source, more than just energy storage [11, 12]. The by-product oxygen has many useful applications ranging from green disinfectant ozone to direct consumption in oxygen-based processes. Although hydrogen production from solar energy has been reported to have very high efficiency, recent literature of NERL reported that the conversion of sunlight into hydrogen has achieved a "world record" of 16.2% via a photo-electrochemical (PEC) water-splitting process [13], which surpassed the previous record of 14% of Fraunhofer ISE and the California Institute of Technology [14, 15]. However, to be a sustainable energy source, the production efficiency must be very high to meet the energy demand. The present study focuses on the utilization of the highly efficient CPV technology for the production of hydrogen as solar energy storage and sustainable future fuel. On the other hand, due to operation under high concentration, there is a lot of heat dissipation by MJC, which can be recovered for low-temperature solar thermal applications.

We propose to have compact and low-cost CPV system which can exploit the scale of mass production methods and can be placed on rooftop or open field applications. The proposed compact and low-cost CPV design approach will eradicate the gigantic design complexities, with market boost by enabling it to tap one of the biggest potentials of rooftop installation by common user, which current commercial CPV systems are not capable to do. Such proposed compact design will be utilized to develop a CPV-hydrogen/oxygen system to meet the sustainable future energy goals using solar energy. The produced hydrogen will be stored for ease, which acts as a portable energy to readily convert it back to electricity using a fuel cell. Oxygen (O2) can be used to produce a disinfectant Ozone (O3) for its oxidation/cleaning potential.

Such system will not only target domestic sector but majorly the industrial sector depending upon their use, by utilizing existing rooftop space with expected sunlight to electricity efficiency of 28%, sunlight to hydrogen efficiency of 18% and sunlight to electricity/thermal efficiency of 71%.

2 Compact CPV Field Configuration

In order to reduce overall hardware requirements of solar trackers, especially the electronic modules needed for conventional astronomical solar tracking, the position of the sun for initial guess of the solar tracker will be obtained by solar geometry model [16–18] based upon the latitude, longitude, date, and time of the tracker location. Azimuth and zenith are the two solar angles which are used to define the position of the sun, as given by Eqs. (1A) and (2).

If $\omega > 0$,

$$\theta_a = 360 - \left[90 + \cos^{-1}\frac{\sin\delta - \sin(90-\theta_z)\sin\phi}{\cos(90-\theta_z)\cos\phi}\right] \tag{1A}$$

If $\omega < 0$,

$$\theta_a = 90 + \cos^{-1}\frac{\sin\delta - \sin(90-\theta_z)\sin\phi}{\cos(90-\theta_z)\cos\phi} \tag{1B}$$

and

$$\theta_z = \cos^{-1}\{\sin\phi\ \sin\delta + \cos\phi\ \cos\delta\ \cos\omega\} \tag{2}$$

where ø, ω, and δ are latitude, hour, and declination angles, respectively.

These solar position coordinates will be distributed to the entire CPV field of slave trackers with low-cost wireless transceivers of Zigbee modules, Fig. 1. In this entire network, the central microcontroller will act as master which will send the coordinates of the position of the sun and the CPV trackers will act as slave to follow the received solar position as per their stage-1 tracking to guess their initial position. Therefore, we can reduce the expensive electronic module requirements of the overall CPV field with reduced cost as only master microcontroller will be equipped with such units to act as a central information transmitter. For communication between Zigbee modules, Eqs. (3) and (4) are used to assign names to certain control box in CPV field, as explained in [19].

$$Cskip = 1 + C_m(L_m - d - 1) \qquad if\ R_m = 1, \qquad for,\ d < L_m \tag{3}$$

$$Cskip = \frac{1 + C_m - R_m - \left(C_m R_m^{(L_m - d - 1)}\right)}{1 - R_m}, \quad otherwise, \quad for,\ d < L_m \tag{4}$$

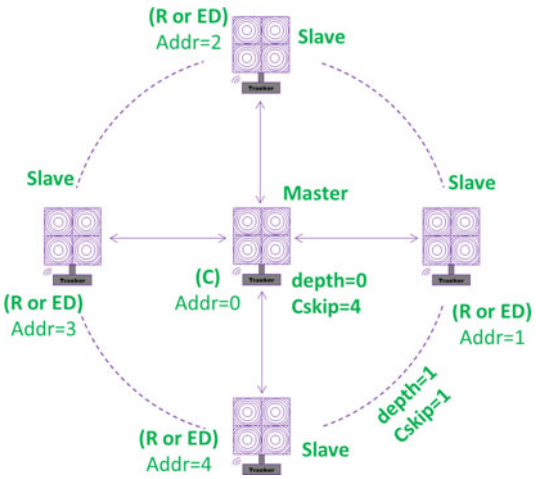


Fig. 1 Master–slave configuration for compact CPV

3 Twin Lens Collimator-Based Solar Tracking Sensor

In order to ensure the tracking accuracy required by the CPV module, i.e., 0.1^{o}, each tracker is equipped with solar feedback sensor, Fig. 2, which accounts for the fine adjustments of the individual tracker position in case of small difference in the solar position and the actual position of the tracker. In order to design a low cost but highly accurate and sensitive solar feedback sensor, an optical collimator with dual lens is used with an array of ordinary photo-sensors. In its simple form, the solar feedback sensor is designed to have a collimated bright spot at the center of the photo-sensor array. With the induced tracking error, the bright spot shifts from the center, and if tracking error exceeds the required tracking accuracy, the solar feedback sensor is aligned with the tracker such that the bright spot hits any of the photo-sensor in the array. The feedback sensor then gives the feedback signal as high, indicating tracking error and the tracker is adjusted such that the bright sport again comes in the center of the photo-sensor array. Another advantage of concentrated bright spot is to hit the photo-sensor with light intensity larger than their saturation limit and in such case, the sensor can only give extreme high or extreme low output. As a result, we can use any ordinary photo-sensor in binary mode by avoiding its nonuniform response under the same light intensity.

4 Prototype of Compact CPV with Hydrogen System and Testing Methodology

By adopting the novel and low-cost solar tracking strategies, the prototype of compact CPV system is developed using triple junction InGaP/InGaAs/Ge solar cell and such a compact CPV system is then be used to build a pilot setup of CPV-hydrogen/oxygen

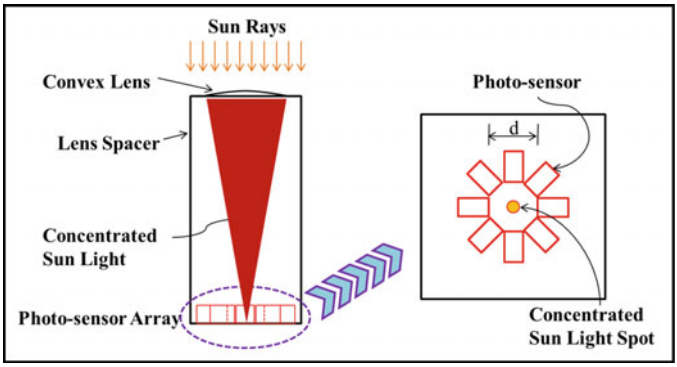


Fig. 2 Twin lens collimator-based solar tracking sensor

system to produce hydrogen and oxygen by electrolytic splitting of water through proton exchange membrane (PEM) electrolyzer, Fig. 3. The main reason to use PEM-based electrolyzer is to have high purity of the gases without any extra energy intensive separation processes. The produced hydrogen and oxygen are compressed to be stored into portable cylinders for further use such as fuel cell or as disinfectant. Such a developed system will be able to be installed anywhere either rooftop or open field, for domestic, commercial, or industrial use, to make solar energy as sustainable energy source with steady power production. The performance of the CPV-hydrogen/oxygen system is analyzed in terms of conversion efficiency as the performance parameter for each component of the system is recorded through central data logging unit.

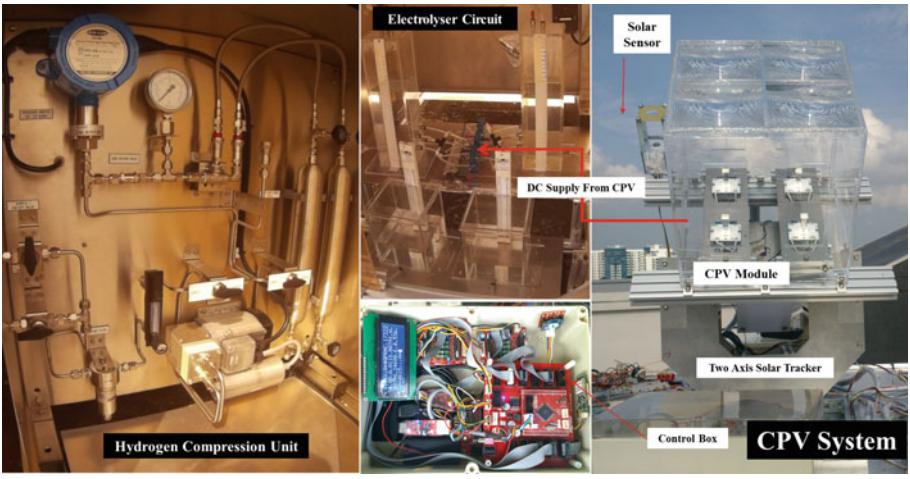


Fig. 3 Prototype of compact CPV with hydrogen system

Based upon the energy received by the CPV system, overall sunlight to hydrogen efficiency can be calculated as

$$\eta_{EL} = \frac{\dot{n}_{E,H2} 237200}{I_E . U_E} = \frac{\eta_{EF} 1.23 N_{EC}}{U_E} \tag{5}$$

$$\eta_{STH} = \frac{P_{out}}{P_{in,solar}} \frac{\left(\dot{n}_{E,H2} 237200\right)}{\left(\frac{P_{CPV}}{\eta_{CPV}}\right)} = \frac{\eta_{EF} I_E 1.23 N_{EC}}{\left(\frac{P_E}{\eta_{CPV}}\right)} = \frac{\eta_{EF} I_E 1.23 N_{EC}}{\left(\frac{U_E I_E}{\eta_{CPV}}\right)} = \eta_{CPV} \frac{\eta_{EF} 1.23 N_{EC}}{U_E} \tag{6}$$

The electrical power output of the CPV system is the same as the electrical energy input of the electrolyzer. The thermal power output of the CPV system is given by Eq. (7).

$$Q_{th} = \left\{ (I_b \times A_{con} \times \eta_{op}) - P_{CPV} \right\} \times (1 - Q_{loss}) \tag{7}$$

Therefore, the efficiency of CPVT system can be calculated as

$$\eta_{CPVT} = \frac{P_{out-CPVT}}{P_{in,solar}} = \frac{P_{CPV} + Q_{th}}{I_b \times A_{con}} \tag{8}$$

In order to calculate the long-term efficiency of the system in form of total energy output against total energy input, the average efficiency of the system can be calculated as [20]

$$\eta_{DNI} = \frac{Electrical\ Rating}{Solar\ DNI\ Input} = \frac{\left[\left(\sum_{i=1}^{t} \left(\frac{\left(V_{CPV} . \mathrm{I}_{CPV}\right)_i - \left(V_{CPV} . \mathrm{I}_{CPV}\right)_{i-1}}{2 \times A_C} \right) \right) \times \frac{365}{m} \right]}{\left[\left(\sum_{i=1}^{t} \left(\frac{\left(\mathrm{I}_b\right)_i - \left(\mathrm{I}_b\right)_{i-1}}{2} \right) \right) \times \frac{365}{m} \right]} \tag{9}$$

5 Results and Discussion

The performance of the developed compact CPV-hydrogen/oxygen system is analyzed for whole day operation with received direct normal irradiance (DNI), as shown in Fig. 4. First, it can be seen that the system showed maximum CPV efficiency of 28%, which is decreasing with a decrease in direct normal irradiance (DNI). The decrease in CPV efficiency is due to the increase in the concentration at cell area, which causes an increase in cell operating temperature. It can be seen that CPVT efficiency of 71% maximum is recorded, which is almost stable throughout the day. With the increase in direct normal irradiance (DNI), as CPV efficiency drops with an increase in cell temperature, the thermal recovery increases. Therefore, the overall efficiency of CPVT system is almost steady throughout the day.

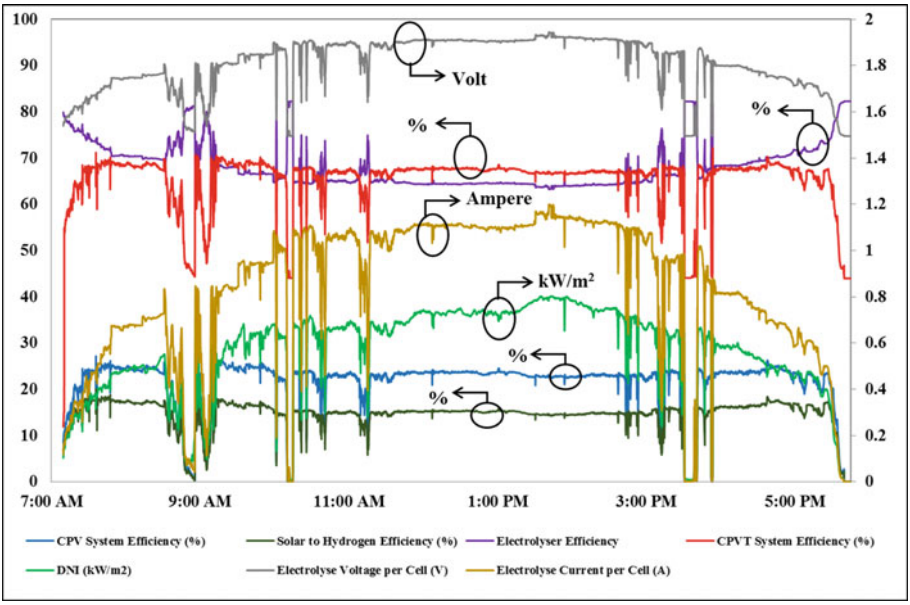


Fig. 4 Performance curves for CPV, CPVT, and CPV-hydrogen system.

For CPV-hydrogen/oxygen, a maximum sunlight to hydrogen (STH) efficiency of 18% is recorded, which decrease with increase in temperature. It can be seen that the decrease in STH efficiency is due to the decrease in the electrolyzer efficiency. From Eq. (5), it is clear that the efficiency of the electrolyzer is inversely proportional to it operating voltage. As the CPV electricity is directly supplied to the electrolyzer, therefore, with an increase in direct normal irradiance (DNI), the power output of CPV increases and so as the operating voltage of electrolyzer. That is why, electrolyzer efficiency is decreasing with an increase in direct normal irradiance (DNI).

Figure 5 shows the long-term field performance of CPV system and compares it with conventional PV systems installed in tropical weather condition of Singapore [21], in terms of its electricity and hydrogen production. The presented data show that CPV has double the capacity of power production than conventional PV system, even in tropical weather conditions which are not considered suitable for CPV operation as it can only respond to beam radiations and tropical weather conditions have less share of beam radiations in received solar energy. In addition, it can also be seen that the CPV production boost to three times the conventional PV output in desert field condition, which is due to larger share of beam radiations in the desert area. However, from hydrogen production point of view, CPV has hydrogen production potential of 5-9 kg/m^2/year which is also two to three times the potential of conventional PV. In the current study, CPV-hydrogen system is presented as a future sustainable energy source with promising production potential and the CPVT system can achieve as high as 71% solar energy conversion efficiency, depending upon its application.

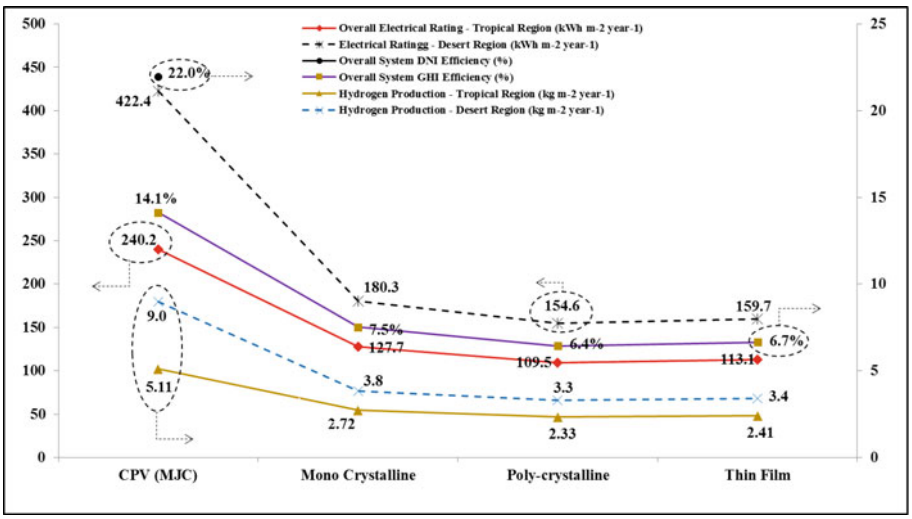


Fig. 5 Long-term electrical rating and performance of CPV and CPV-hydrogen system

6 Conclusion

A design of compact CPV system is presented with in-house built solar tracker sensor and master–slave field configuration, to achieve overall less system cost without compromising the accuracy of the system. The system is designed to handle the installation-related limitations of conventional gigantic (big size) CPV, which will boost the CPV installations in the rooftop of buildings like conventional PV. In addition, overall 71% solar energy efficiency, 28% electrical efficiency and 18% hydrogen production efficiency shows huge performance potential of CPV than conventional PV. For CPVT application with 71% efficiency, although the whole energy cannot be utilized in high grade energy applications, it can be properly utilized with effective heat recovery system applications. With such huge potential and the energy efficiency, concentrated photovoltaic (CPV) can provide a sustainable solar energy system for future green environment. In addition, such compact CPV system will boost the urban rooftop installation of CPV and so as its market share. This study also proves that the feasibility and potential of CPV is the tropical weather conditions.

References

1. Shafiei, E., Davidsdottir, B., Leaver, J., Stefansson, H., Asgeirsson, E.I.: Comparative analysis of hydrogen, biofuels and electricity transitional pathways to sustainable transport in a renewable-based energy system. Energy **83**, 614–627 (2015)
2. Burhan, M., Shahzad, M.W., Ng, K.C.: Development of performance model and optimization strategy for standalone operation of CPV-hydrogen system utilizing multi-junction solar cell. Int. J. Hydrog. Energy **42**(43), 26789–26803 (2017)

3. Burhan, M., Chua, K.J.E., Ng, K.C.: Electrical rating of concentrated photovoltaic (CPV) systems: long-term performance analysis and comparison to conventional PV systems. Int. J. Technol. **7**(2), 189–196 (2016). https://doi.org/10.14716/ijtech.v7i2.2983
4. Ng, K.C., Burhan, M., Shahzad, M.W., Ismail, A.B.: A universal isotherm model to capture adsorption uptake and energy distribution of porous heterogeneous Surface. Sci. Rep. **7**(1), 10634 (2017)
5. Burhan, M., Chua, K.J.E., Ng, K.C.: Sunlight to hydrogen conversion: design optimization and energy management of concentrated photovoltaic (CPV-Hydrogen) system using micro genetic algorithm. Energy **99**, 115–128 (2016)
6. IPCC (2012): Renewable energy sources and climate change mitigation. special report of the intergovernmental panel on climate change
7. Burhan, M., Chua, K.J.E., Ng, K.C.: Simulation and development of a multi-leg homogeniser concentrating assembly for concentrated photovoltaic (CPV) system with electrical rating analysis. Energy Convers. Manag. **116**, 58–71 (2016)
8. Burhan, M., Shahzad, M.W., Ng, K.C.: Long-term performance potential of concentrated photovoltaic (CPV) systems. Energy Convers. Manag. **148**, 90–99 (2017)
9. Burhan, M., Oh, S.J., Chua, K.J.E., Ng, K.C.: Double lens collimator solar feedback sensor and master slave configuration: Development of compact and low cost two axis solar tracking system for CPV applications. Solar Energy **137**, 352–363 (2016)
10. Burhan, M., Chua, K.J.E., Ng, K.C.: Long term hydrogen production potential of concentrated photovoltaic (CPV) system in tropical weather of Singapore. Int. J. Hydrog. Energy **41**(38), 16729–16742 (2016)
11. Burhan, M., Shahzad, M.W., Choon, N.K.: Hydrogen at the rooftop: compact CPV-hydrogen system to convert sunlight to hydrogen. Appl. Therm. Eng. **132**, 154–164 (2018)
12. Burhan, M., Oh, S.J,, Chua, K.J., Ng, K.C.: Solar to hydrogen: Compact and cost effective CPV field for rooftop operation and hydrogen production. Appl. Energy 2016 (article in press). http://dx.doi.org/10.1016/j.apenergy.2016.11.062
13. Arndt, R., Puto, I.R.: Basic understanding of IEC standard testing for photovoltaic panels. TUV SUD America Inc. http://www.tuvamerica.com/services/photovoltaics/articlebasicunderstandingpv.pdf. Accessed 21 Oct 2015
14. Burhan, M., Shahzad, M.W., Ng, K.C.: Sustainable cooling with hybrid concentrated photovoltaic thermal (CPVT) system and hydrogen energy storage. Int. J. Comput. Phys. Ser. **1**(2), 40–51 (2018)
15. Burhan, M., Shahzad, M.W., Oh, S.J., Ng, K.C.: A pathway for sustainable conversion of sunlight to hydrogen using proposed compact CPV system. Energy Convers. Manag. **165**, 102–112 (2018)
16. Muhammad, B., Seung, J.O., Ng, K.C., Chun, W.: Experimental investigation of multijunction solar cell using two axis solar tracker. In: Applied Mechanics and Materials, vol. 819, pp. 536–540. Trans Tech Publications (2016)
17. Oh, S.J., Burhan, M., Ng, K.C., Kim, Y., Chun, W.: Development and performance analysis of a two-axis solar tracker for concentrated photovoltaics. Int. J. Energy Res. **39**(7), 965–976 (2015)
18. Astronomical Applications Department of the U.S. Naval Observatory. http://aa.usno.navy.mil/data/docs/AltAz.php. Accessed 10 Apr 2015

19. Farahani, S.: ZigBee wireless networks and transceivers. Newnes, 2011. In: International Energy Agency, CO_2 Emissions from Fuel Combustion Highlights (2013)
20. Burhan, M.: Theoretical and experimental study of concentrated photovoltaic (CPV) system with hydrogen production as energy storage (Doctoral dissertation) (2015)
21. Lin-Heng, L., Lok-Ming, C., Chia, A., Gunawansa, A., Harn-Wei, K.: Sustainability Matters, Challenges and Opportunities in Environmental Management in Asia. Pearson Education South Asia Pte Ltd., Singapore (2011)

The Influence of Square Wire Attack Angle on the Heat Convection from a Surrogate PV Panel

Yang Yang[1(✉)], Ashhar Ahmed[1], David S-K. Ting[1], and Steve Ray[2]

[1] Turbulence and Energy Laboratory, University of Windsor, 401 Sunset Ave, Windsor, ON, Canada
yangl9n@uwindsor.ca
[2] Essex Energy, Oldcastle, ON, Canada

Abstract. Passively enhancing the convective cooling of solar photovoltaics (*PV*) panels by the prevailing atmospheric wind is possibly the simplest and yet an effective means to boost solar energy production. Among the numerous ways to promote turbulent convective cooling, the conventional tripwire concept is straightforward to implement. The square wire with sharp edges placed at varying angles of attack is explored for its effectiveness in augmenting the convective cooling of a surrogate *PV* panel. Specifically, the heat transfer performance over an approximately constant temperature flat plate downstream of a 4 mm^2 wire, placed at 6 mm from the flat surface, with 15° to 75° attack angle is scrutinized in a wind tunnel at 5 m/s wind velocity. The resulting local Nusselt number enhancement is explained in terms of the perturbed turbulent flow characteristics, detailed using a triple-wire constant-temperature hot wire. A higher level of turbulence along with more energetic and larger energy-containing eddying motions are captured behind the 60° attack angle wire, contributing to the most efficacious cooling of the surrogate *PV* panel. These results translate into 2–3% increase in the *PV* panel energy output.

Nomenclature

A	Area of the *PTFE* plate
D	Width of the square wire
E	Uncertainty
F	Frequency
G_T	Solar irradiance
h	Heat transfer coefficient
K_{air}	Thermal conductivity of air
K_{PTFE}	Thermal conductivity of *PTFE* plate
Re	Reynolds number
Nu	Nusselt number
Nu_0	Nusselt number without the square wire
P	Electric power output
P_0	Electric power output without the square wire
$PTFE$	Polytetrafluoroethylene
PV	Solar photovoltaics

A. Vasel and D. S.-K. Ting (Eds.): EAS 2018, SPE, pp. 103–128, 2019.
https://doi.org/10.1007/978-3-030-00105-6_7

$Q_{convection}$	Convective heat transfer
$Q_{radiation}$	Radiation heat transfer
Q_{total}	Total heat transfer
St	Strouhal Number
T_{air}	Temperature of the ambient air
T_c	Cell temperature of *PV* panel
T_{bottom}	Temperature of the bottom surface of the *PTFE* plate
T_{ref}	Reference temperature
T_{top}	Temperature of the top surface of the *PTFE* plate
$T_{top,0}$	Temperature of the top surface of the *PTFE* plate without the wire
T_{wall}	Wall temperature of the wind tunnel
t_{PTFE}	Thickness of the *PTFE* plate
Tu, Tv, Tw	Turbulence intensity in X, Y, and Z direction, respectively
U, V, W	Instantaneous velocity in X, Y, and Z direction, respectively
$\bar{U}, \bar{V}, \bar{W}$	Time-averaged velocity in X, Y, and Z direction, respectively
U_∞	Free stream velocity in X direction
u, v, w	Instantaneous fluctuating velocity in X, Y, and Z direction, respectively
$u_{rms}, v_{rms}, w_{rms}$	Root mean square velocity in X, Y, and Z direction, respectively
X	Streamwise direction
Y	Widthwise direction
Z	Vertical direction
β	Temperature coefficient
ε	Emissivity
η	Electric efficiency
η_{ref}	Efficiency at reference temperature
Θ	Square wire attack angle
Λ	Integral length scale
λ	Taylor microscale
ν	Kinematic viscosity
τ	Temporal distance
τ_Λ	Integral time scale
τ_λ	Taylor time scale

1 Introduction

Global energy needs will expand by 30% from 2017 to 2040. Among worldwide end-uses of energy, the demand of the electricity keeps rising, making up 40% of the increase in final consumption to 2040 [1]. To meet the rising demand of the electricity and the challenge of decarbonising power supply, the global investment in electricity has already overtaken that of oil and gas in 2016. Among others, due to the rapid development of solar photovoltaics (*PV*), solar energy has the potential to become the

largest source of low-carbon capacity by 2040, by which time the share of all renewables in total power generation will reach 40% [1].

The power output efficiency of *PV* panels is influenced by solar irradiance and their cell temperature. For *PV* panels with polycrystalline modules, the efficiency can drop to 9.1% if cell temperature is above 60 °C [2]. Thus, much effort has been invested in finding better ways to keep *PV* panels cool. The numerous attempts can be divided into two groups; active versus passive cooling approaches. Drawing water across the back of the *PV* panel is an effective means. For example, Baloch et al. [3] investigated the effect of converging angle of water cooling channel on *PV* cooling and concluded that the two-degree angle was observed to have the best thermal performance, which reduced cell temperature from 71 °C to 45 °C, and improved power output by 35.5% in the month of June. An experimental study of a novel *PV* cooling technique with water spray was conducted by Nizetic et al. [4]. Their results showed that it was possible to achieve an increase of 16.3% in electric power output on a summer day, by reducing the *PV* panel temperature from 52 °C up to 24 °C. Moreover, Bhattacharjee et al. [5] investigated the back surface cooling approach with different designs of heat-absorbing pipe (semi-oval serpentine, circular spiral, and circular spiral semi-flattened). They found that the circular spiral-shaped semi-flattened copper pipe showed the best performance and helped the maximum power of the panel increase by 16.8% (from 22.6 to 26.39 W). However, the active techniques mentioned above require external power.

Due to the considerations of overall power conservation and cost, many passive *PV*-cooling techniques have also widely been explored. For example, Rajput and Yang [6] investigated the cooling effect of a clear coated aluminum heat sink. It was determined by their study that natural convection was present from the heat sink facing downwards, reducing temperature by 25 °C given an 88.1 °C *PV* panel. Also promising is the passive technique of enhancing convective heat transfer, as turbulent generators have drawn great attention in recent years. For example, Fouladi et al. [7] investigated the turbulent flow behind a finite-height perforated plate and concluded that the turbulent flow was not in its continuous decay mode with downstream distance due to the interference of the flow over the upper edge of the perforated plate, resulting in an increase of turbulence intensities with downstream distance at some points, which is helpful to enhance convective heat transfer. Subsequently, they explored the effect of the triangular rib and the square rib on a flat plate [8, 9]. They found that the rib caused a larger boundary layer and a higher level of turbulence intensity in the area close to the wall, which are predicted to affect the heat transfer from the surface. Moreover, Wu at el focused on the studies about the turbulent flow behind a delta winglet [10, 11], and illustrated that vortices were observed downstream of the winglet, resulting in a higher level of turbulence, which is expected to have a higher heat transfer rate than that associated with the base flat plate case.

Among different type of turbulence generators, the studies about the ribs were widely developed due to their structural simplicity. Marumo et al. [12] investigated the turbulent heat transfer in a flat plate behind a cylinder and found that the gap between the cylinder and the plate was helpful for heat transfer in the near wake. Wang and Sundén [13] studied the heat transfer behind the truncated ribs, and concluded that the vortices generated from the end effect was able to sweep cooler fluid towards the wall, resulting in a better heat transfer in the close-rib region. In another study, the rib shape

effect on heat transfer was explored by Kamali and Binesh [14], which illustrated the trapezoidal shaped rib had a better heat transfer performance than others. Subsequently, an experimental investigation associated with the effect of the trapezoidal rib chamfering angle on the heat transfer was conducted by Ali et al. [15], which points out the changing angle had an influence on the location of separation and reattachment, and the resulting augmentation of heat transfer. Besides, the arrangements of the rib also draw a great deal of attention recently. For example, Liu et al. [16] explored the heat transfer using different arrangements of the square ribs, and concluded that the flow path become more complex by the staggered arrangements, leading to a larger heat transfer enhancement. A numerical investigation of heat transfer in two-pass ribbed channels was developed by Ravi et al. [17]. And they found that the V-shaped-arranged ribs provided the maximum heat transfer enhancement in the first and second pass, while the M-shaped-arranged ribs provided the best thermal performance in the bend region. It is clear from the sample of studies that heat transfer performance behind a rib is significantly influenced by its geometry, operating parameters and arrangements. In spite of great efforts, much remains to be discovered about the heat transfer behind the rib. The current study aims at furthering the state-of-the-art knowledge by focusing on the promising influence of the attack angle of a square wire on the heat transfer from a flat surface.

2 Experimentation

The experiment was carried out in a closed-loop wind tunnel with a 760 mm^2 cross-section. A 295 mm by 380 mm polytetrafluoroethylene (*PTFE*) plate was inlaid in the center of the test section base, which is made of 10 mm thick fiberglass with thermal conductivity of 0.04 Wm^{-1} K^{-1}. The 3 mm thick *PTFE* plate has a thermal conductivity of 0.25 Wm^{-1} K^{-1} and an emissivity of 0.92 [18, 19]. A water tank was heated up to generate steam, keeping the bottom surface of the *PTFE* plate at a uniform and constant temperature of 100 °C. A FLIR C2 thermal camera was used to capture the temperature distribution of the top surface. The thermal image had 60 by 80 pixels, resulting in a spatial resolution of approximately 4.5 mm (Fig. 1).

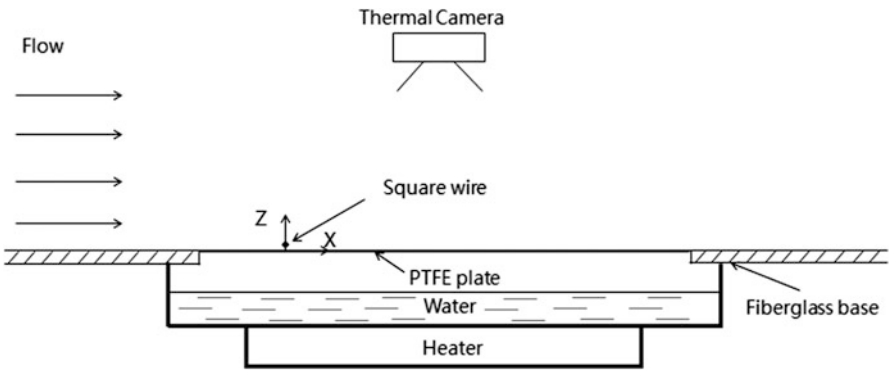


Fig. 1. Heat transfer experimental setup.

A 4 mm^2 wire that spanned 630 mm (more than twice *PTFE* plate width) was used to perturb the flow with respect to the reference, empty wind tunnel without the wire case; see Fig. 2. The square wire was placed at 6 mm above the *PTFE* plate, resulting in a 1.5D gap. The attack angle was varied from 15 to 75° in 15° increments, while the wind speed was fixed at approximately 5 m/s. X, Y (into the paper) and Z in Fig. 2 represent the streamwise, widthwise, and vertical direction, respectively. The point of origin, x = 0, corresponds to the leading edge of the lower part of the square wire.

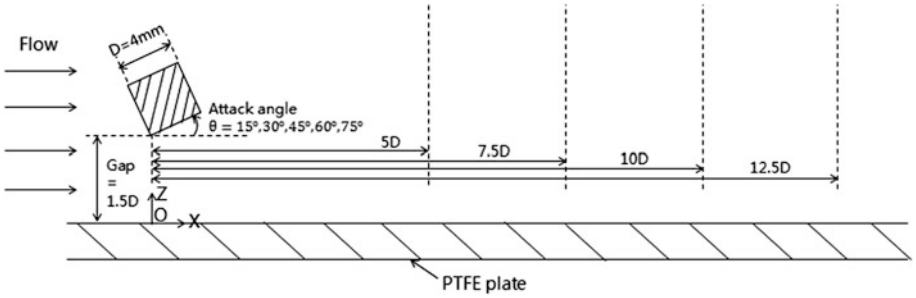


Fig. 2. Square wire and flow measurement locations.

A triple sensor hotwire probe (type 55P95) with a constant temperature anemometer was employed to obtain the velocities at 5D, 7.5D, 10D and 12.5D downstream of the square wire; see Fig. 2. Each measured line was from 4 mm (1D) to 40 mm (10D) above the plate surface, with a spatial resolution of 1 mm. The velocity signals were low passed at 30 kHz, and 10^6 measurements were sampled at 80 kHz.

3 Data Processes

3.1 Heat Transfer Enhancement

The total heat transferred away from the *PTFE* surface can be deduced from:

$$Q_{total} = K_{PTFE} A \left(\frac{T_{bottom} - T_{top}}{t_{PTFE}} \right) \tag{1}$$

Here, K_{PTFE} is the thermal conductivity of the *PTFE*, which is 0.25 Wm^{-1} K^{-1}, A is the area of the *PTFE* surface, t_{PTFE} is the thickness of the *PTFE* plate, T_{bottom} and T_{top} are the temperatures at the bottom and top surfaces of the *PTFE* plate, respectively. As the conduction to the rest of the wind tunnel fiberglass base is negligible, the total heat transfer consists of radiation and convection. The convective heat transfer, key parameter of this study, can be obtained from:

$$Q_{convection} = Q_{total} - Q_{radiation} \tag{2}$$

The radiation heat transfer can be estimated from:

$$Q_{radiation} = \varepsilon\sigma A\left(T_{top}^4 - T_{wall}^4\right) \tag{3}$$

where Boltzmann's constant, σ is 5.67×10^{-8} Wm^{-2} K^{-4} [20], emissivity ε is 0.92 and T_{wall} is the wall temperature of the wind tunnel.

The convective heat transfer coefficient can be calculated:

$$h = \frac{Q_{convection}}{A\left(T_{top} - T_{air}\right)} \tag{4}$$

The corresponding non-dimensional Nusselt number, representing the ratio of convective heat transfer to conductive heat transfer, can be obtained from:

$$Nu = \frac{hL}{K_{air}} \tag{5}$$

where L is the characteristic length and K_{air} is the thermal conductivity of air. As the focus is on convective heat transfer enhancement, the normalized Nusselt number is introduced:

$$\frac{Nu}{Nu_0} = \frac{h}{h_0} \tag{6}$$

where Nu_0 and h_0 are the Nusselt number and the convective heat transfer coefficient for the unperturbed (smooth) surface case without the square wire. It is worth noting the advantages of adopting this normalized Nusselt number. First, it does not depend on the characteristic length. Secondly, as long as the circulating wind tunnel air is at the same thermodynamic state, K_{air} is of no concern.

3.2 Predicted Panel Power Output Increase

The electric efficiency of the *PV* panel is defined as:

$$\eta = \frac{P}{AG_T} \tag{7}$$

where P is electric power output, G_T is the solar irradiance. This efficiency is affected by cell temperature and solar irradiance, and can be simplified as [2]:

$$\eta = \eta_{ref}\left[1 - \beta\left(T_c - T_{ref}\right)\right] \tag{8}$$

where η_{ref} is the efficiency at the reference temperature, β is the temperature coefficient, T_c is the cell temperature, and T_{ref} is the reference temperature (typically 25 °C). Combining the two equations on efficiency, the power output can be expressed as:

$$P = AG_T\eta_{ref}\left[1 - \beta\left(T_c - T_{ref}\right)\right] \tag{9}$$

In this study, we assume that the cell temperature, T_c, is equal to the temperature of the top surface of the *PTFE* plate, T_{top}. Thus, the power output increase compared with the un-wired case can be expressed as:

$$\frac{P}{P_0} \times 100\% = \left[\frac{1 - \beta\left(T_{top} - T_{ref}\right)}{1 - \beta\left(T_{top,0} - T_{ref}\right)}\right] \times 100\% \tag{10}$$

where we assume $\beta = 0.43\%$ in this study [21], and P_0 is power output of un-wired case.

3.3 Flow Characteristics

With the 3D hotwire, the instantaneous velocities in the X, Y and Z directions were measured simultaneously. Each of these three components consists of the corresponding time-averaged velocity plus the fluctuating part, i.e.,

$$U = \bar{U} + u; V = \bar{V} + v; W = \bar{W} + w \tag{11}$$

where $\bar{U}, \bar{V}, \bar{W}$ are the respective time-averaged velocities, and u, v, w, the corresponding instantaneous fluctuating velocities. The time-averaged velocity can be calculated from:

$$\bar{U} = \frac{1}{N}\sum_{i=1}^{N} U_i; \bar{V} = \frac{1}{N}\sum_{i=1}^{N} V_i; \bar{W} = \frac{1}{N}\sum_{i=1}^{N} W_i \tag{12}$$

where $N = 10^6$ is the sampling number.

The root-mean-square velocity,

$$u_{rms} = \sqrt{\sum_{i=1}^{N} \frac{u_i^2}{N-1}}; v_{rms} = \sqrt{\sum_{i=1}^{N} \frac{v_i^2}{N-1}}; w_{rms} = \sqrt{\sum_{i=1}^{N} \frac{w_i^2}{N-1}} \tag{13}$$

This fluctuating velocity can be normalized by the freestream velocity to give the relative turbulence intensity,

$$Tu = \frac{u_{rms}}{U_\infty}; Tv = \frac{v_{rms}}{U_\infty}; Tw = \frac{w_{rms}}{U_\infty} \tag{14}$$

Some form of vortex shedding is expected from the square wire under certain conditions. To describe the vortex shedding frequency, the normalized frequency, Strouhal number is introduced:

$$St = \frac{fD}{U_\infty} \tag{15}$$

Other than velocity fluctuation, turbulence length scale is another key parameter to describe the flow field turbulence [22]. The integral scale, representing the energy containing large eddies, can be derived from the autocorrelation function of turbulence fluctuating velocities. The autocorrelation function in the temporal distance is:

$$f(\tau) = \frac{\overline{u(t)u(t+\tau)}}{\overline{u^2(t)}} \tag{16}$$

This can be converted into the corresponding spatial autocorrelation function based on Taylor frozen hypothesis [23],

$$f(r) = f(\tau)\bar{U} \tag{17}$$

The integral time scale is defined as:

$$\tau_\Lambda = \int_0^\infty f(\tau)d\tau \tag{18}$$

The integration process was terminated at the first zero crossing point. From this, invoking the Taylor frozen hypothesis [10] gives

$$\Lambda = \bar{U}\tau_\Lambda \tag{19}$$

At the other end of the turbulence energy cascade, the Taylor microscale associated with the dissipation is

$$\lambda = \bar{U}\tau_\lambda \tag{20}$$

The Taylor timescale,

$$\tau_\lambda = \sqrt{\frac{\overline{2u^2}}{\left(\frac{du}{dt}\right)^2}} \tag{21}$$

3.4 Uncertainty Analysis

The total uncertainty (E) of the measurement parameters consists of the Bias and the Precision:

$$E = \sqrt{\text{Bias}^2 + \text{Precision}^2} \tag{22}$$

where Precision follows the Student's t distribution method with a 95% confidence interval. With sample measurements repeated 10 times, $t = 2.262$.

The uncertainties of the other dependent parameters were determined following the propagation rule [24]. For example, if $y = f(x_1, x_1, \ldots, x_n)$, the uncertainty of y can be deduced from:

$$E(y) = \sqrt{\sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i} E(x_1)\right]^2} \tag{23}$$

The total uncertainty of the temperature of the top surface of the *PTFE* plate (T_{top}) can be calculated from Eq. (22). The precision was deduced from a sample of 10 repeats for the wire with a 60° attack angle as it represents the average case. The uncertainty of each parameter associated with heat transfer can be deduced by the propagation of the uncertainty; Eq. (23):

$$E(Q_{total}) = \sqrt{\left[\frac{\partial Q_{total} E(T_{top})}{\partial T_{top}}\right]^2} = \frac{K_{PTFE} A}{t_{PTFE}} E(T_{top}) \tag{24}$$

$$E(Q_{radiation}) = \sqrt{\left[\frac{\partial Q_{radiation} E(T_{top})}{\partial T_{top}}\right]^2} = 4\varepsilon\sigma A T_{top}^3 E(T_{top}) \tag{25}$$

$$\begin{aligned} E(Q_{convection}) &= \sqrt{\left[\frac{\partial Q_{convection} E(Q_{total})}{\partial Q_{total}}\right]^2 + \left[\frac{\partial Q_{convection} E(Q_{radiation})}{\partial Q_{radiation}}\right]^2} \\ &= \sqrt{[E(Q_{total})]^2 + [E(Q_{radiation})]^2} \end{aligned} \tag{26}$$

$$\begin{aligned} E(h) &= \sqrt{\left[\frac{\partial h E(Q_{convection})}{\partial Q_{convection}}\right]^2 + \left[\frac{\partial h E(T_{top})}{\partial T_{top}}\right]^2} \\ &= \sqrt{\left[\frac{E(Q_{convection})}{A(T_{top} - T_{air})}\right]^2 + \left[\frac{Q_{convection} E(T_{top})}{A(T_{top} - T_{air})^2}\right]^2} \end{aligned} \tag{27}$$

$$E\left(\frac{Nu}{Nu_0}\right) = \sqrt{\left[\frac{\partial \frac{Nu}{Nu_0} E(h)}{\partial h}\right]^2 + \left[\frac{\partial \frac{Nu}{Nu_0} E(h_0)}{\partial h_0}\right]^2} = \sqrt{\left[\frac{E(h)}{h_0}\right]^2 + \left[\frac{h E(h_0)}{h_0^2}\right]^2} \tag{28}$$

$$E\left(\frac{P}{P_0}\right) = \sqrt{\left[\frac{\partial \frac{P}{P_0} E(T_{top})}{\partial T_{top}}\right]^2 + \left[\frac{\partial \frac{P}{P_0} E(T_{top,0})}{\partial T_{top,0}}\right]^2}$$
$$= \sqrt{\left[\frac{-\beta E(T_{top})}{1-\beta(T_{top,0}-T_{ref})}\right]^2 + \left[\frac{-\beta[1-\beta(T_{top}-T_{ref})]E(T_{top,0})}{[1-\beta(T_{top,0}-T_{ref})]^2}\right]^2} \tag{29}$$

For the hotwire measurement, the Bias uncertainty of $\bar{U}$ and u_{rms} can be estimated from the error of calibration [25], while the Precision was deduced from the 10-time repeatability test, which took place at the point X = 5D, Z = 1D behind the square wire with an attack angle of 60°. According to Eq. (22), the total uncertainty $E(\bar{U})$ and $E(u_{rms})$ were obtained as summarized in Table 1.

Table 1. Typical uncertainties of mean and rms velocity

Parameter	$\bar{U}$	u_{rms}	$\bar{W}$	w_{rms}
Uncertainty	0.101 ms^{-1}	0.045 ms^{-1}	0.103 ms^{-1}	0.036 ms^{-1}

Then, the uncertainty of the normalized time-averaged and rms velocity can be calculated:

$$E(\bar{U}/U_\infty) = \frac{\bar{U}}{U_\infty}\sqrt{2\left[\frac{E(\bar{U})}{\bar{U}}\right]^2} \tag{30}$$

$$E(u_{rms}/U_\infty) = \frac{u_{rms}}{U_\infty}\sqrt{\left[\frac{E(\bar{U})}{\bar{U}}\right]^2 + \left[\frac{E(u_{rms})}{u_{rms}}\right]^2} \tag{31}$$

Following the propagation rule:

$$E(u) = \sqrt{E(U)^2 + E(\bar{U})^2} \tag{32}$$

$$E(\overline{uw}) = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\left[w_i^2 E^2(u_i) + u_i^2 E^2(w_i)\right]} \tag{33}$$

$$E(\overline{uw}/U_\infty^2) = \frac{\overline{uw}}{U_\infty^2}\sqrt{\left[\frac{E(\overline{uw})}{\overline{uw}}\right]^2 + \left[\frac{2E(\bar{U})}{\bar{U}}\right]^2} \tag{34}$$

According to Refs [10] and [26] and the propagation of error, the uncertainty of the integral scale and Taylor microscale can be expressed as:

$$E(\Lambda) = \Lambda\sqrt{\left[\frac{E(\bar{U})}{\bar{U}}\right]^2 + 2\left[\frac{2E(u_{rms})}{u_{rms}}\right]^2} \tag{35}$$

$$E(\lambda) = \lambda\sqrt{\left[\frac{E(\bar{U})}{\bar{U}}\right]^2 + 2\left[\frac{E(u_{rms})}{u_{rms}}\right]^2} \tag{36}$$

$$E\left(\frac{\Lambda}{D}\right) = \frac{\Lambda}{D}\sqrt{\left[\frac{E(\Lambda)}{\Lambda}\right]^2 + \left[\frac{E(D)}{D}\right]^2} \tag{37}$$

$$E\left(\frac{\lambda}{D}\right) = \frac{\lambda}{D}\sqrt{\left[\frac{E(\lambda)}{\lambda}\right]^2 + \left[\frac{E(D)}{D}\right]^2} \tag{38}$$

where the uncertainty of D was estimated from 10 repeated measurements using a ruler with a resolution of 0.5 mm. This resulted in $E(D)$ = 0.261 mm.

Representative uncertainties, as described above, were performed. Sample uncertainty values of the studied parameters are provided in Table 2.

Table 2. Representative uncertainties of studied parameters.

Parameter	T_{top}	Nu/Nu_0	P/P_0	U/U_∞	W/U_∞
Uncertainty	2.02 °C	0.21	0.016	0.030	0.022
Parameter	u_{rms}/U_∞	w_{rms}/U_∞	$\overline{uw}/U_\infty^2$	Λ/D	λ/D
Uncertainty	0.012	0.010	0.001	0.142	0.004

4 Results and Discussion

In this section, the heat transfer results are discussed in the context of *PV* panel power output enhancement first. Subsequently, the physics behind the varying localized heat convection improvement up to 50D downstream is explained in terms of the detailed flow characteristics induced by the square wire.

4.1 Heat Transfer and Predicted Power Output Enhancement

The effect of the square wire attack angle on the convective heat transfer performance of the flat plate is depicted in Fig. 3. The results are portrayed in terms of the normalized Nusselt number, illustrating the convective heat transfer enhancement with respect to the un-wired case. It is clear from the figure that, the normalized Nusselt number increases with streamwise distance in the near-wire region, and it reaches a peak at around 5 $\sim$ 8D downstream of the square wire. Farther downstream, the heat transfer augmentation slowly drops, indicating the flow is restored to a less perturbed condition. Moreover, the higher Nu/Nu_0 value can be observed behind the 60° attack

angle wire, which indicates better convective heat transfer performance than that associated with the other attack angles.

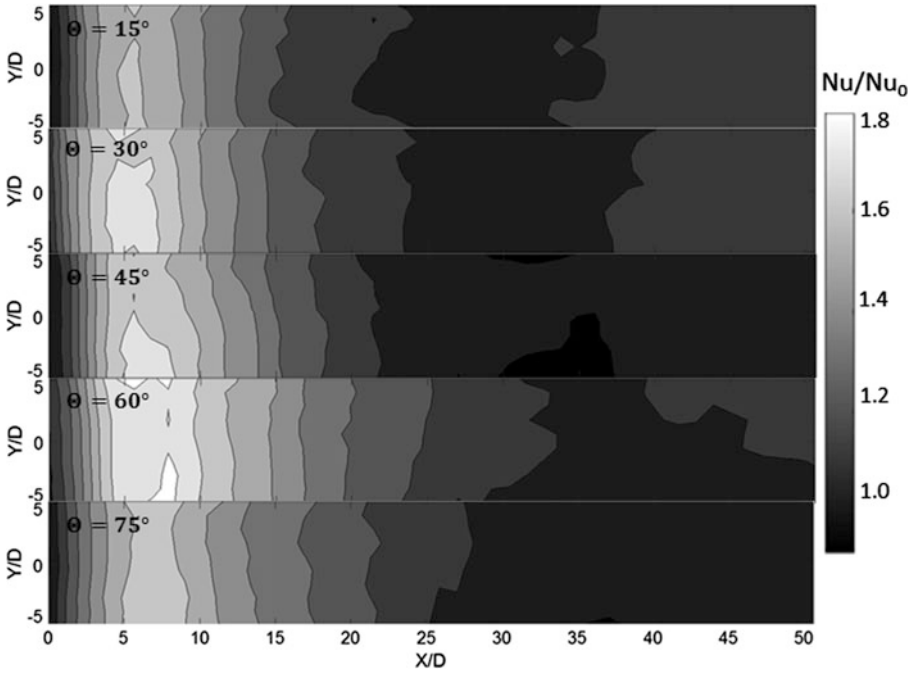


Fig. 3. Contour of normalized Nusselt number (Nu/Nu_0) downstream of the square wire.

As the studied case is strictly two dimensional, with some unavoidable inhomogeneity in the Y direction, averaging the results across the ±5D in the Y direction can better delineate the variation in the convective heat transfer improvement in the streamwise direction. Figure 4 shows this span-averaged normalized Nusselt number downstream of the square wire. The peak values can be clearly found at 5–8D downstream, varied with respect to the attack angle. The square wire with an attack angle of 60° gives the highest peak value of Nu/Nu_0, around 1.8, at about 8D downstream. This superior performance of the 60° attack angle wire continued all the way until roughly 26D downstream, before the 15° case out-performed it marginally. It is interesting to note that, over the studied surface area, this 15° turbulence generator promotes the least cooling. To better understand the heat transfer trend, the flow characteristics are examined next.

Figure 5 shows the area-averaged (area encompassed by Y = − 5 to 5D, and Z = 0 to 20 or 50D) temperature and predicted power output increase of the surrogate *PV* panel, of which the temperature coefficient (β) is assumed to be 0.43% (Clear Power CS6P-230P, poly-crystalline [21]). It is clear from the figure that, over the entire studied area, the surrogate *PV* panel performs best when the attack angle is 60°, for the studied conditions. In this case, the area-averaged temperature from X/D = 0 to 50

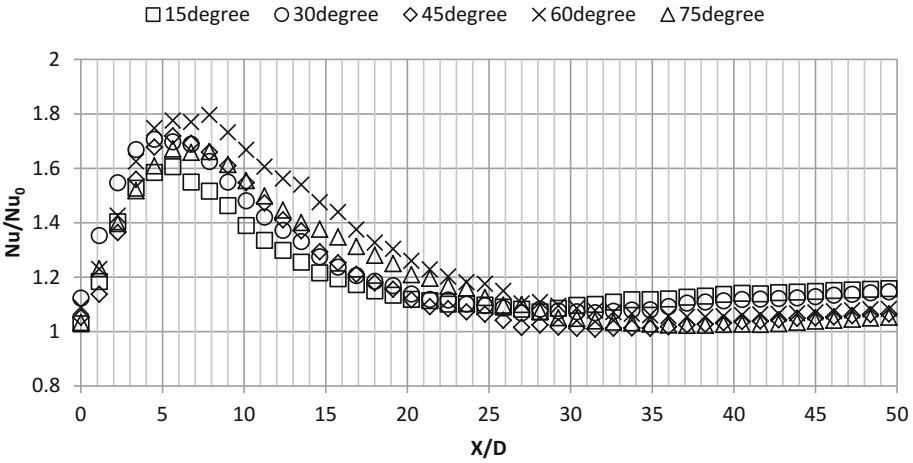
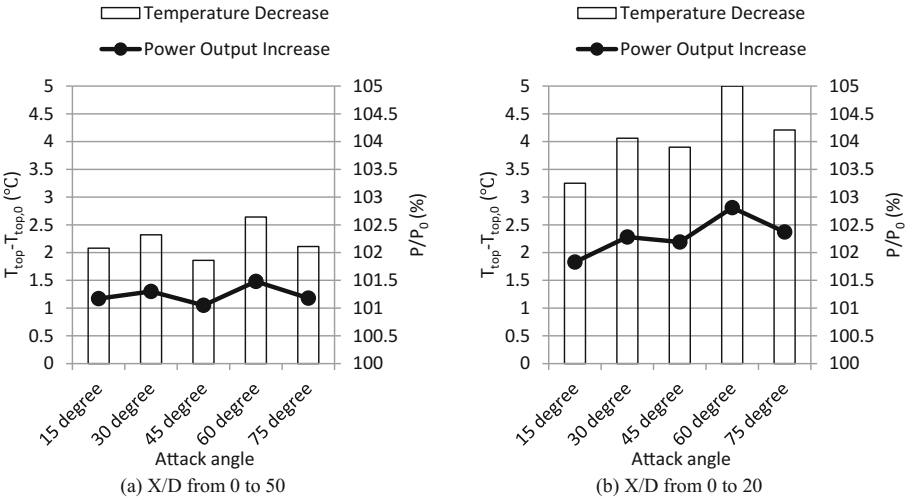


Fig. 4. Span-averaged normalized Nusselt number downstream of the square wire.

decreases by 2.6 °C (from 80.09 °C to 77.45 °C), compared with the un-wired flat plate. This leads to a maximum of 1.5% power output increase for the 60° attack angle wire. If we focus on the near wake region, from X/D = 0 to 20, the area-averaged temperature of the *PV* panel behind the 60° square wire is approximately 5 °C cooler than that without the wire. The practical effect is a significant power output boost of up to 2.8%. In order to explain the mechanism behind the difference of heat transfer improvement, the flow characteristics generated by the square wire will be detailed in the next section.



(a) X/D from 0 to 50 (b) X/D from 0 to 20

Fig. 5. The area-averaged temperature and the predicted power output increase.

4.2 Flow Characteristics

4.2.1 Velocity Profile

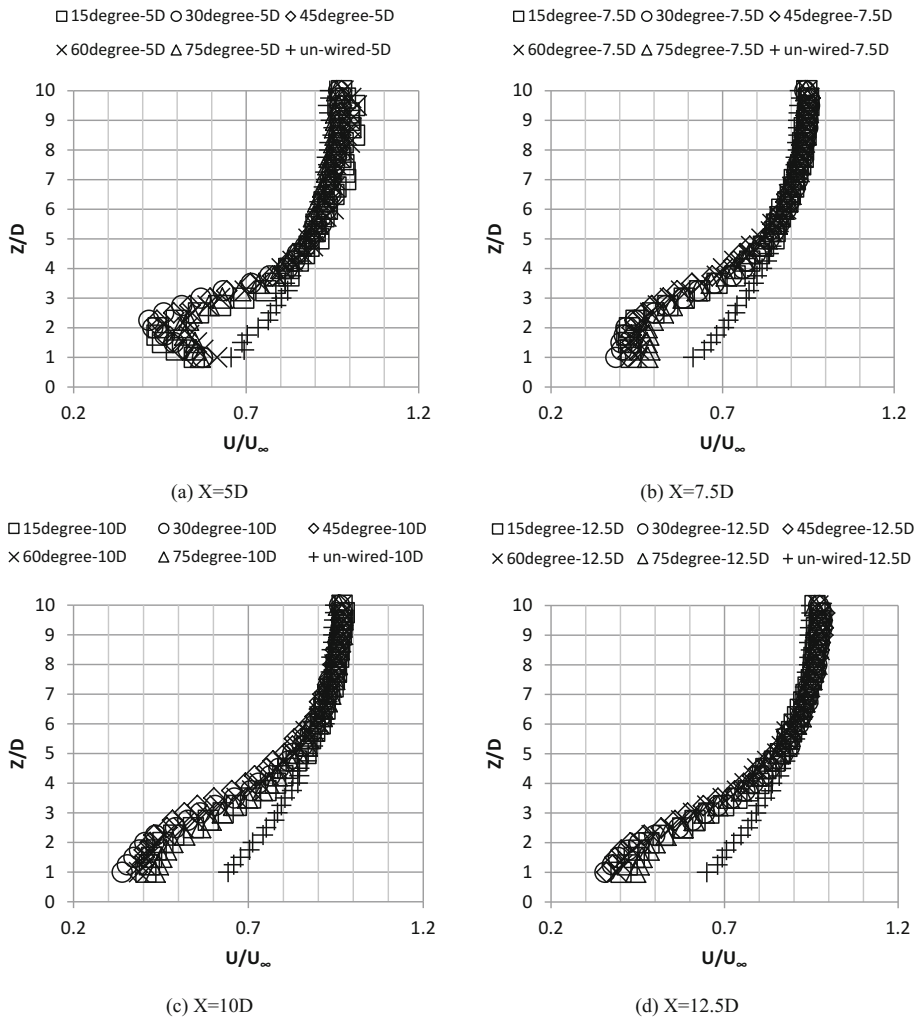
Figure 6 shows the time-averaged velocity profile in the X direction behind the square wire at different attack angle. It is clear from the figure that at 5D downstream of the square wire, velocity has a high value at a height of 1D from the plate, regardless of the attack angle. This is caused by the jet flow coming from the gap between the wire and the plate. Farther away from the plate, in the increasing Z direction, the velocity profile experiences a deficit at $Z \approx 2D$. This signifies the wake region behind the wire. From the experimental study of Dutta et al. [27], the drag coefficient were affected by the orientation of the square cylinder, and the minimum drag coefficient was observed at the attack angle of 22.5°, (same as 67.5° due to the symmetry), as well as the minimum velocity deficit. On the other hand, Dayem and Bayomi [28] found that the minimum drag coefficient occurred at the square cylinder with attack angle of 30° (same as 60°). In the current study, the best performing 60° attack angle wire induced the least amount of velocity deficit; other than the un-wired reference case. The first runner-up, the 75° attack angle wire (see Fig. 5), seems to prevail on this farther downstream. It should be stressed that as far as the local heat transfer is concern, the near-plate (or wall) velocity is most important. Namely, a larger local velocity would more effectively sweep the thermal energy out of the way. The overall velocity profile indicates, among other things, the amount of shearing (and thus, turbulence) and how the flow will further evolve. It is also clear that the jet emerging from the 1.5D gap between the wire and the plate does not persist far downstream.

Other than the streamwise velocity for this largely 2-dimensional case, the velocity in the Z direction is also of importance. As such, the velocity profile in the Z direction is plotted in Fig. 7. We can see from the graphs that, at 5D behind the wire, the W value starts from a large positive value (velocity vector pointing upward, away from the plate) at the closest-to-the-plate measurement points at $Z/D \approx 1$; fear of hot-wire breakage prevented us from venturing closer). This upward, away from the plate, flow is good for convecting heat away. Notwithstanding, it is imperative to note that the magnitude of W is an order of magnitude small than that of U; compare Fig. 7 with Fig. 6.

The switching from a positive W to a negative one with increasing Z in Fig. 7 implies flow circulation, or, more correctly, it signifies a recirculation region. Physically, this is also beneficial as far as heat convection is concerned. The two most effective 'coolers' ($\theta = 60°$ and 75°) do not distinguish themselves with a stronger recirculation. This is most likely due to the aforementioned small W magnitude, limited by the omnipresent experimental uncertainties. While some negative, downward velocity persists beyond $X \approx 12.5D$, the positive, upward W deteriorates rapidly, and is no longer seen shortly after $X \approx 5D$. This appears to be another confirmation that the jet through the gap between the wire and the plate is weak.

4.2.2 Turbulence Intensity

In order to further understand the flow structure and its effect on heat transfer, turbulence intensity is detailed in this section. Figure 8 shows the streamwise turbulence



(a) X=5D (b) X=7.5D

(c) X=10D (d) X=12.5D

Fig. 6. U/U_∞ profile downstream of the square wire.

intensity profile. The general trend of decreasing turbulence with distance downstream of the wire is obvious. The perturbed and significantly thickened boundary layer, on the other hand, persisted far downstream. Future study should aim at bringing the persisting high-*Tu* flow closer to the plate. At 5D downstream, two peaks corresponding to the two shear layers from the upper and bottom edges of the square wire. When the flow reaches 7.5D downstream, these *Tu* peaks are no longer obvious, due to the interaction between the shear layers from the bottom edge of the wire and the wall boundary layer. At 10D and 12.5D downstream, the turbulence intensity close to the wall (Z = 1D) drops to approximately 0.1, (compared to around 0.8 in the un-wired case), which presumably indicates the decrease of the cooling effect. Figure 8(a) shows

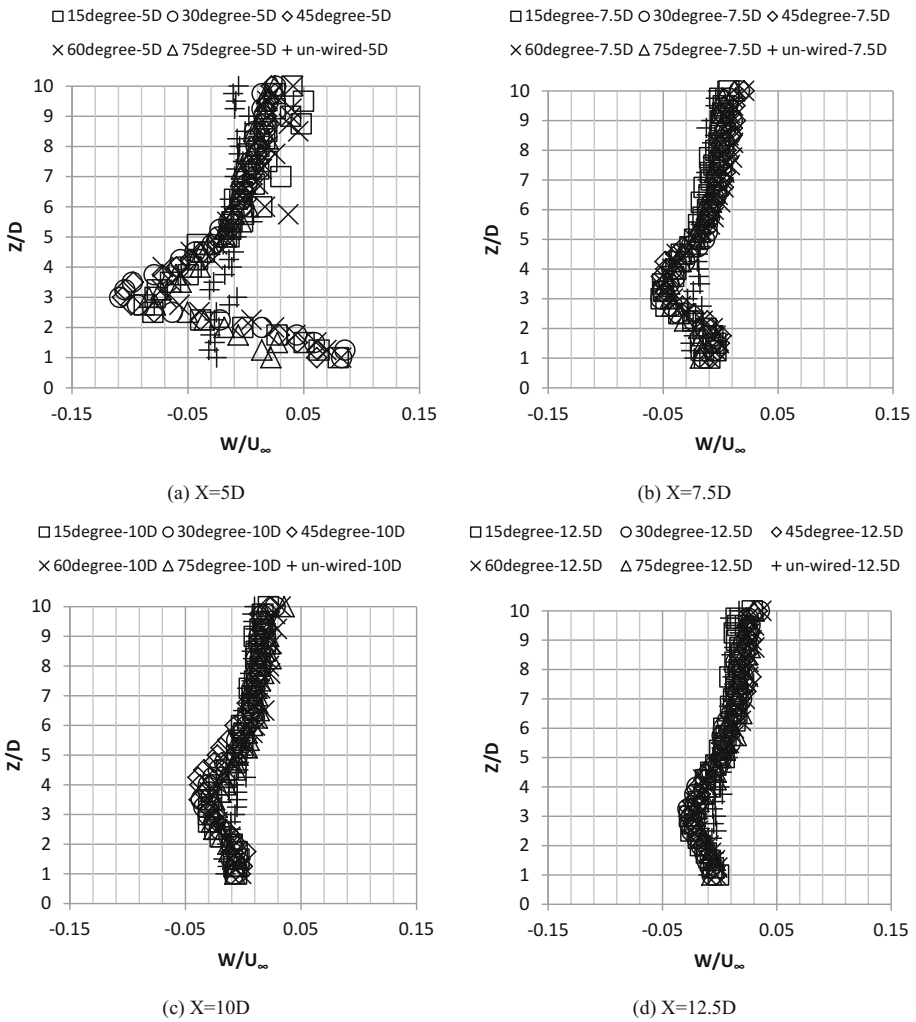


Fig. 7. W/U_∞ profile downstream of the square wire.

the comparison of the *Tu* value at 5D downstream among the different studied attack angles. It is clear that, the turbulence intensity behind the square wire with an attack angle of 60° has a higher value than others, which translates to a better normalized Nusselt number, discussed earlier.

Profiles of the vertical (Z component) turbulence intensity (*Tw*) are depicted in Fig. 9. A peak value can be observed clearly at 3D above the plate, presumably generated by the shear layer from the upper edge of the square wire. With increasing distance downstream, the *Tw* value decreases, indicating the decay of the wire-induced turbulence. Among others, the square wires with attack angles of 60° and 75° have the

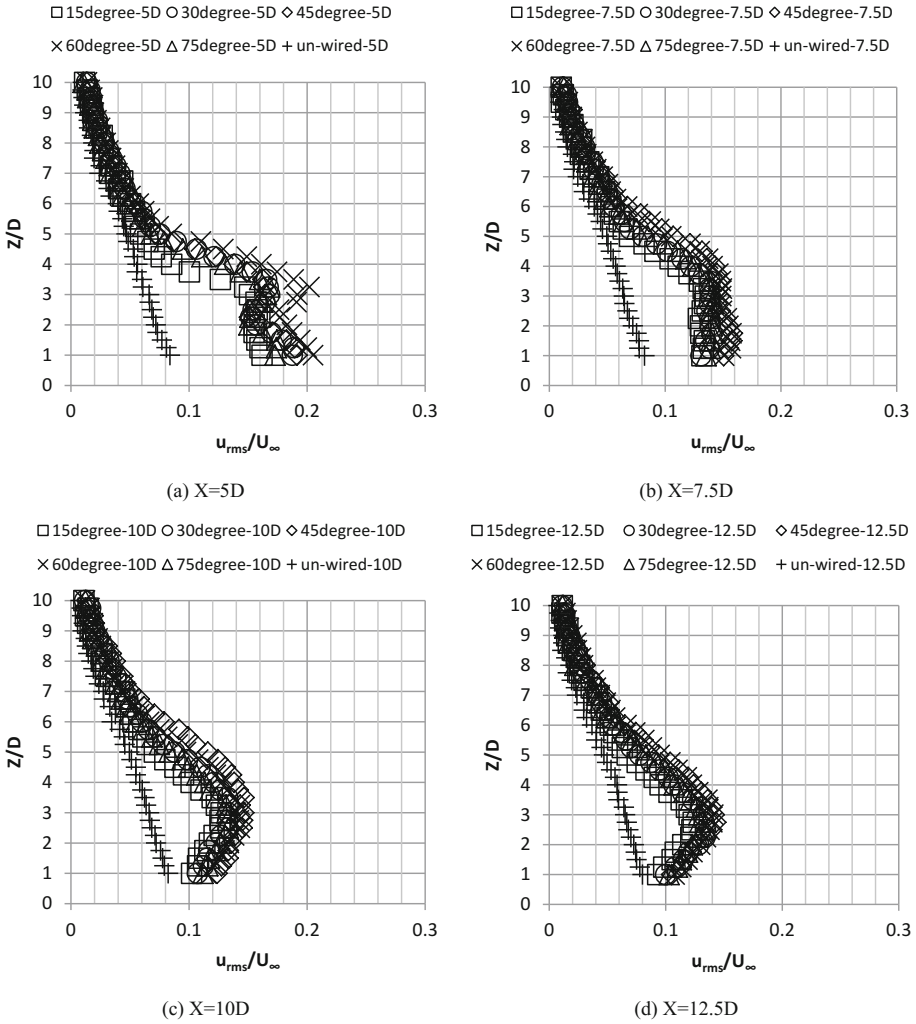


Fig. 8. *Tu* profile downstream of the square wire.

largest *Tw* values, while the 15° wire has the smallest one. These outcomes further support the respective largest and lowest heat transfer enhancement results.

4.2.3 Power Spectrum

In addition to the turbulence intensity, the power spectrum of the fluctuating velocity is also of significance. As a bluff body is utilized, some regular vortex shedding is expected under certain conditions. The resulting vortices can play a significant role in convection heat transfer. The kinetic energy of the flow at 5D downstream of the square wire with an attack angle of 45°, as expressed in power spectrum is depicted in Fig. 10.

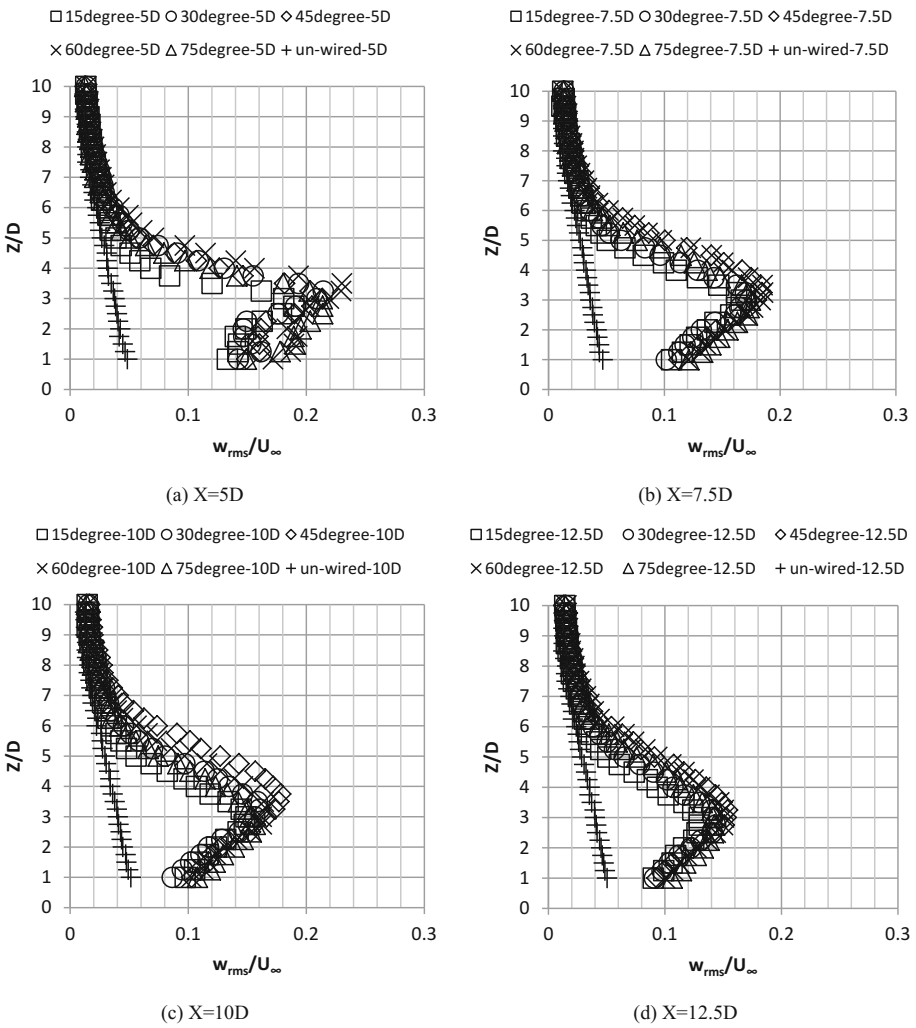
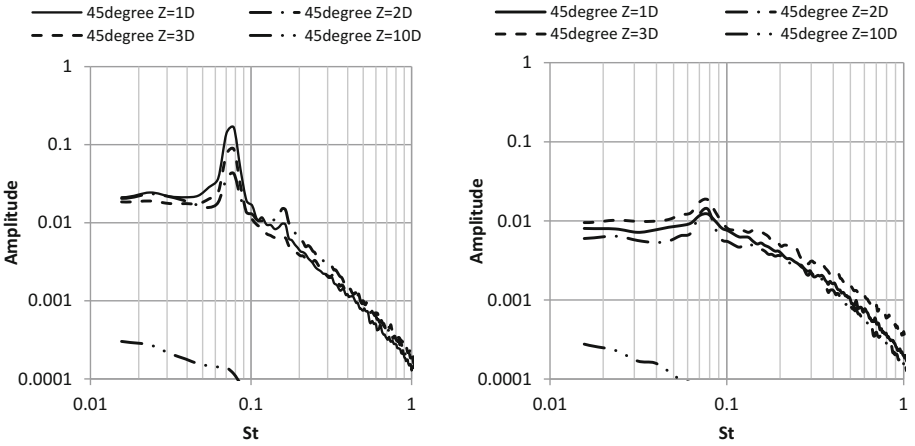


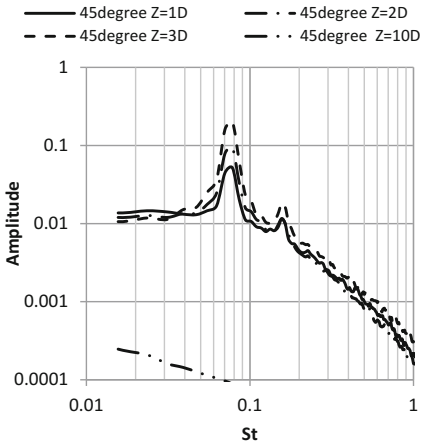
Fig. 9. Vertical turbulence intensity, *Tw*, profile downstream of the square wire.

The Y component power spectrum does not have an obvious peak, while the X and Z components supply a distinct a peak at a Strouhal number of approximately 0.08. This implies that the vortex generated by the square wire is primarily in the XZ-plane. Even with a significant attack angle and close to the plate, the obtained $St = 0.08$ is close to $St = 0.12$ a normal square cylinder [29], at the same Reynolds number, $Re = U_\infty D/v \approx 1270$). Note that the ground effect has been found to inhibit wake development, and thus, reduce the *St* value [30, 31]. Also notable is the occurrence of the most powerful X component vortex shedding at approximately 1D above the plate, while that of the Z component is at around 3D, at which the turbulence intensity reaches the highest value. A possible explanation for this phenomenon is that the flow



(a) X component

(b) Y component



(c) Z component

Fig. 10. Power spectrum at 5D behind the square wire with an attack angle of 45° (Z = 1/2/3/10D)

volume at Z = 1D is not only affected by the layer from the bottom edge of the wire, but also the wall boundary layer; see Fig. 11. The interaction of these two fluctuating turbulent shear layers tend to enhance the X fluctuating component while weakening the Z component.

Figure 12 portrays the comparison of the power spectrum between the square wire with varying attack angles. The largest peak amplitude value occurs for the 60° case, while the 15° wire resulted in the smallest peak value. These results coincide with the results of the turbulence intensity, which indicates that the 60° wire produces a higher

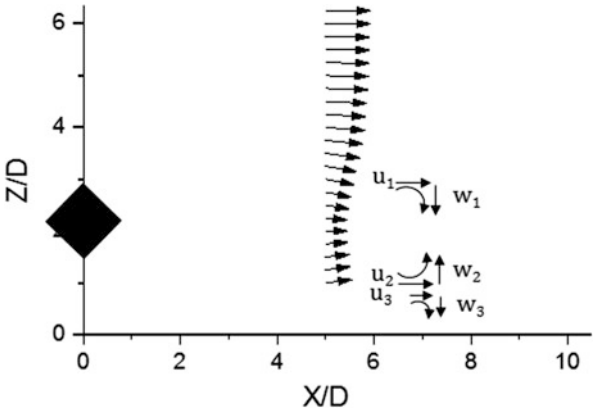


Fig. 11. Velocity vector at 5D behind the square wire and the schematic diagram of the stress

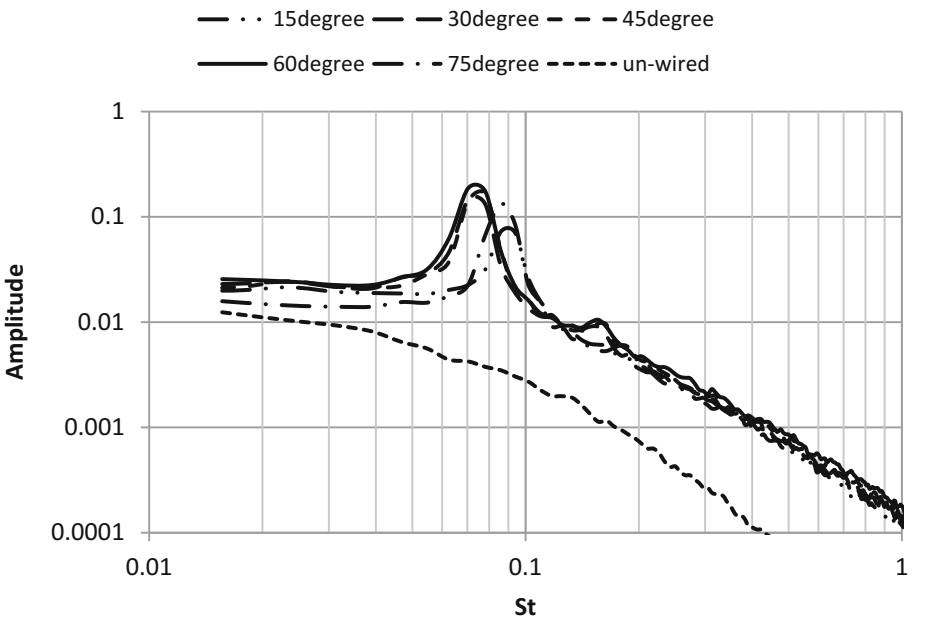
level of turbulence, or, more correctly, the strongest vortex shedding, leading to the most effective heat convection. It is worth emphasizing that the location of the vortex shedding is equally important. What is desirable is a street of vortices fluctuating up and down close enough to the heated surface, to shed the hot fluid near the surface into the freestream, and the cooler fluid unto the surface.

4.2.4 Reynolds Stress

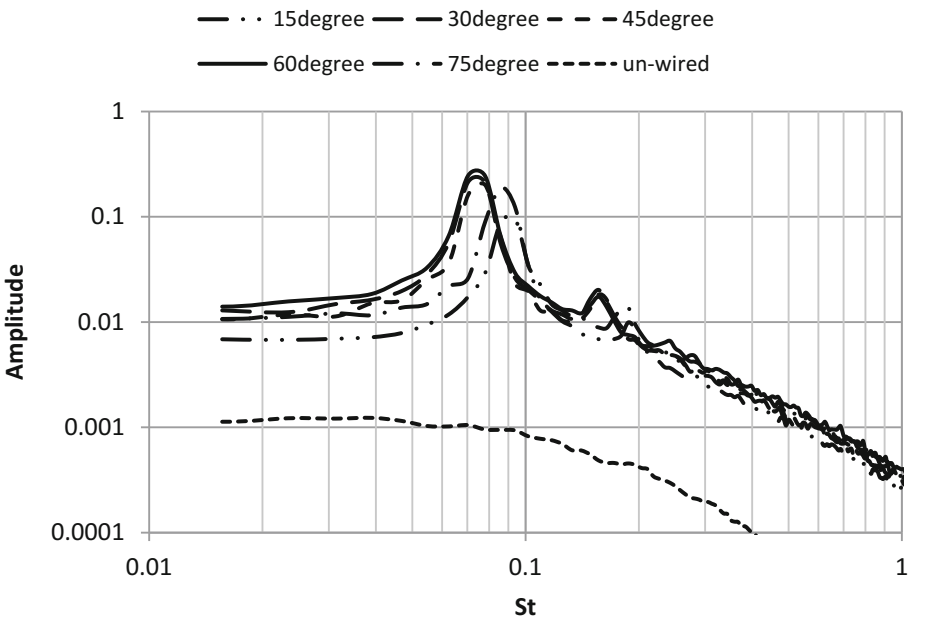
Figure 13 shows the Reynolds stress to complement the illustration of the flow structures such as regular vortices and bulk flow shears. The Reynolds stress is associated with the spatial gradients in time-averaged velocity, it connotes the wall friction and thus, convective heat transfer [32]. The maximum negative Reynolds stress occurs at Z $\approx$ 3D, where the peak values of turbulence intensity in the X and Z directions occur. Also, at 5D and 7.5D downstream of the wire, very large positive Reynolds stress can be observed near the plate, presumably due to the jet flow generated from the gap. Farther downstream, this jet generated $\overline{uw}$ decays and quickly approaches the unwired case at Z = 1D. From Fig. 13(a), it is also clear that, the absolute values of the Reynolds stress behind the 60° wire are higher than that associated with the other attack angles.

4.2.5 Integral Scale

The size of the eddying motion with respect to the thermal boundary layer, to a large extent, decides how effective the shoveling of the thermal energy from the heated panel is. As such, the streamwise integral scale for the region where the turbulence is of significant intensity ($Tu > 0.1$) is depicted in Fig. 14. The enlarging integral scale above a height of approximately 2D is closely tied to the diminishing turbulence level, see Fig. 8. Shortly downstream of the wire, at X = 5D, the integral scale also increases below Y = 2D; see Fig. 14(a). This increase in the integral length probably has more to



(a) X component (Z=1D)



(b) Z component (Z=3D)

Fig. 12. The effect of attack angle on the power spectra at 5D behind the square wire.

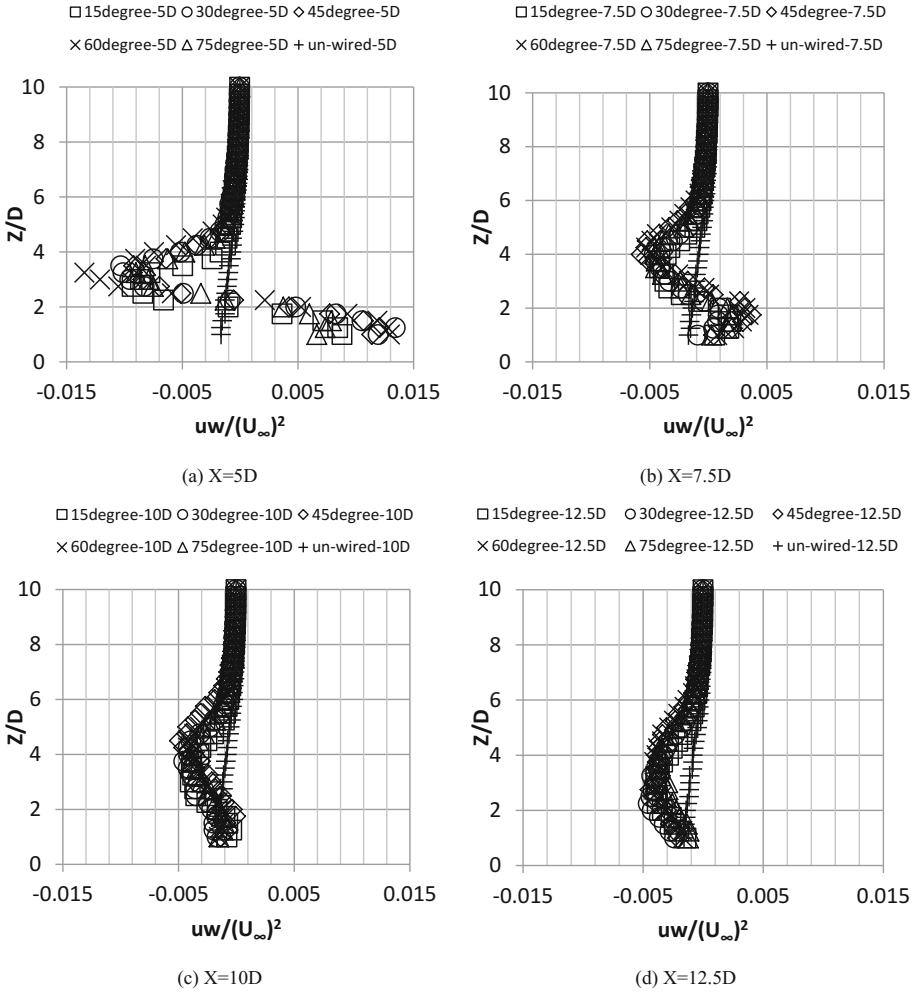


Fig. 13. Normalized reynolds stress profile in the XZ plane.

do with the jet ensuing from the gap between the wire and the plate. In other words, due to the shear layers generated by the upper and lower edges of square wire, the eddies at around 1D and 3D are somewhat larger). Presumably due to interference of the regular vortices shed from the cylinder, the integral scale results at some locations, especially farther downstream, are not very conclusive.

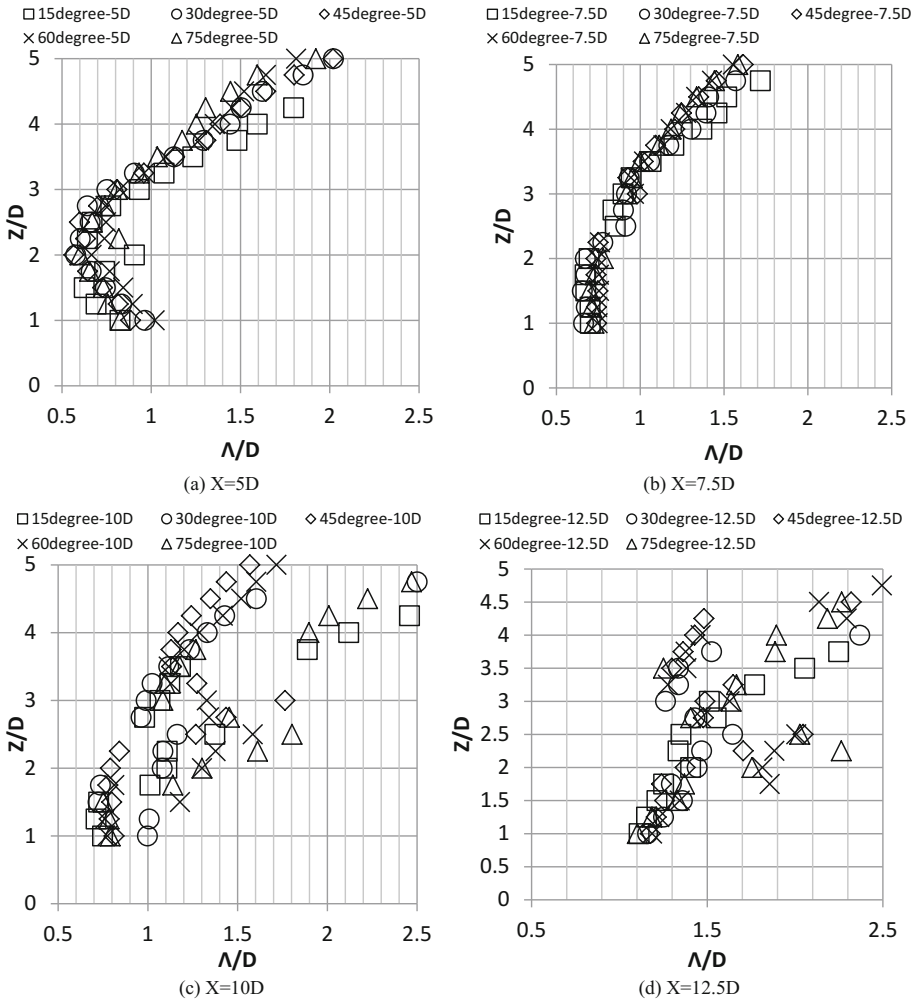


Fig. 14. The streamwise integral scale profile.

4.2.6 Taylor Microscale

Figure 15 depicts the size of the dissipative Taylor microscale. It is clear from the figure that the smallest Taylor microscale values are detected Z ≈ 2D, which corresponds nicely to the more intense turbulence region. The general increase in the Taylor microscale with distance downstream is consistent with the turbulence decay.

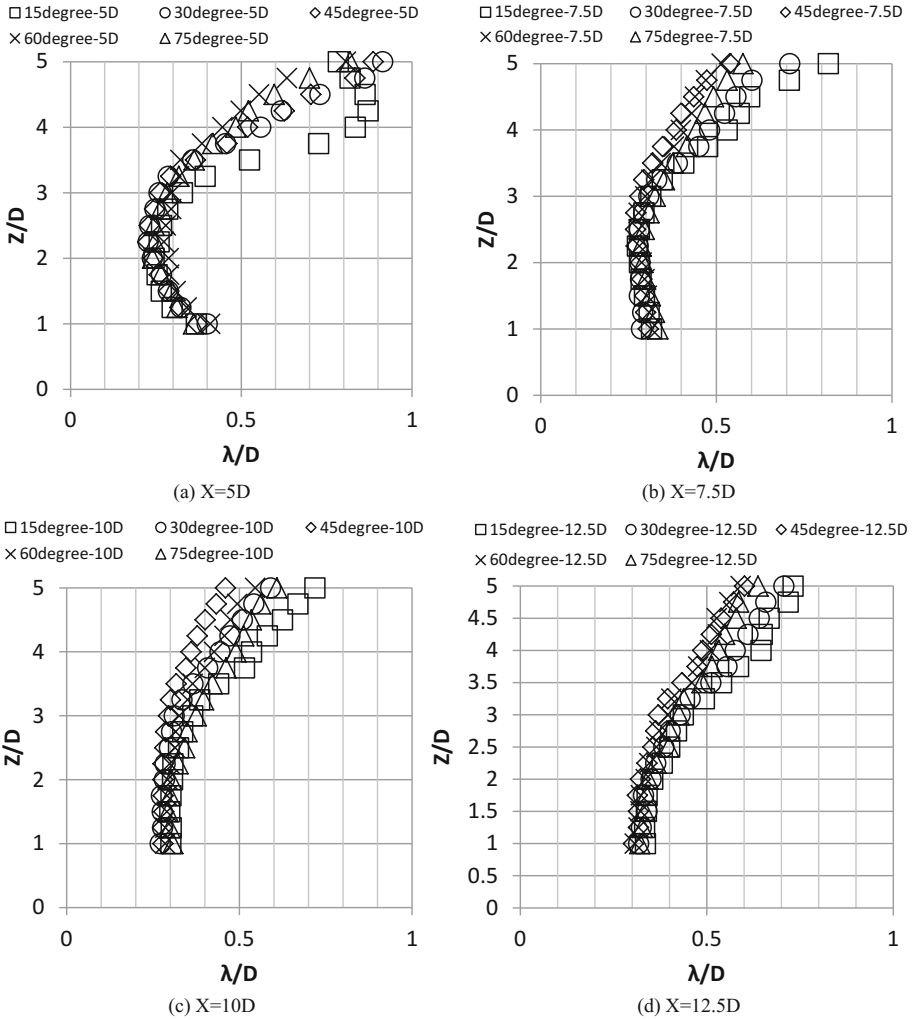


Fig. 15. The streamwise Taylor microscale profile.

5 Conclusions

A D = 4 mm^2 wire attack has been employed to scrutinize the effect of its attack angle on the effectiveness of convecting heat away from a surrogate *PV* panel at 5 m/s wind. A 1.5D gap between the wire and the plate allowed a jet to be emanated, in addition to generating a highly turbulent flow with vortex streets. The Nusselt number normalized by the base unperturbed (by the wire) Nusselt number, Nu/Nu_o, has a peak value at around 5 to 8D downstream. The square wire with an attack angle of 60° promoted the highest heat transfer enhancement, with a maximum Nu/Nu_o value of around 1.8. This

corresponds to about 2.8% power output increase for the *PV* panel. Intense turbulence with large energy-containing eddies, and somewhat broad-banded vortex shedding, appear to furnish the best flow environment for convection heat transfer augmentation.

Acknowledgements. This work was made possible by Natural Sciences and Engineering Research Council of Canada and Ontario Centres of Excellence.

References

1. International Energy Agency: World Energy Outlook 2017 – Executive Summary (2017)
2. Evans, D.L.: Simplified method for predicting photovoltaic array output. Sol. Energy **27**(6), 555–560 (1981)
3. Baloch, A.A.B., Bahaidarah, H.M.S., Gandhidasan, P., Al-Sulaiman, F.A.: Experimental and numerical performance analysis of a converging channel heat exchanger for PV cooling. Energy Convers. Manag. **103**, 14–27 (2015)
4. Nizetic, S., Coko, D., Yadav, A., Grubisic-Cabo, F.: Water spray cooling technique applied on a photovoltaic panel: The performance response. Energy Convers. Manag. **108**, 287–296 (2016)
5. Bhattacharjee, S., Acharya, S., Potar, A., Meena, A., Bairwa, D.S., Meena, D., Bhagora, M., Meena, P., Gautam, P.K.: An investigational back surface cooling approach with different designs of heat-absorbing pipe for PV/T system. Int. J. Energy Res. **42**(5), 1921–1933 (2018)
6. Rajput, U.J., Yang, J.: Comparison of heat sink and water type PV/T collector for polycrystalline photovoltaic panel cooling. Renew. Energy **116**, 479–491 (2018)
7. Fouladi, F., Henshaw, P., Ting, D.S.-K.: Turbulent flow over a flat plate downstream of a finite height perforated plate. J. Fluids Eng. **137**(2), 021203–021212 (2014)
8. Fouladi, F., Henshaw, P., Ting, D.S.-K.: Effect of a triangular rib on a flat plate boundary layer. J. Fluids Eng. **138**(1), 011101–011111 (2015)
9. Fouladi, F., Henshaw, P., Ting, D.S.-K., Ray, S.: Flat plate convection heat transfer enhancement via a square rib. Int. J. Heat Mass Transf. **104**, 1202–1216 (2013)
10. Wu, H., Ting, D.S.-K., Ray, S.: An experimental study of turbulent flow behind a delta winglet. Exp. Thermal Fluid Sci. **88**, 46–54 (2017)
11. Wu, H., Ting, D.S.-K., Ray, S.: The effect of delta winglet attack angle on the heat transfer performance of a flat surface. Int. J. Heat Mass Transf. **120**, 117–126 (2018)
12. Marumo, E., Suzuki, K., Sato, T.: Turbulent heat transfer in a flat plate boundary layer disturbed by a cylinder. Int. J. Heat Fluid Flow **6**(4), 241–248 (1985)
13. Wang, L., Sundén, B.: Experimental investigation of local heat transfer in a square duct with continuous and truncated ribs. Exp. Heat Transf. **18**, 179–197 (2005)
14. Kamali, R., Binesh, A.R.: The importance of rib shape effects on the local heat transfer and flow friction characteristics of square ducts with ribbed internal surfaces. Int. Commun. Heat Mass Transf. **35**, 1032–1040 (2008)
15. Ali, M., Tariq, A., Gandhi, B.: Flow and heat transfer investigation behind trapezoidal rib using PIV and LCT measurements. Exp. Fluids **54**, 1520 (2013)
16. Liu, J., Hussain, S., Wang, J., Wang, L., Xie, G., Sundén, B.: Heat transfer enhancement and turbulent flow in a high aspect ratio channel (4:1) with ribs of various truncation types and arrangements. Int. J. Therm. Sci. **123**, 99–116 (2018)
17. Ravi, B.V., Singh, P., Ekkad, S.V.: Numerical investigation of turbulent flow and heat transfer in two-pass ribbed channels. Int. J. Therm. Sci. **112**, 31–43 (2017)

18. Thermal Conductivity of common Materials and Gases. https://www.engineeringtoolbox.com/thermal-conductivity-d_429.html. Accessed 15 June 2018
19. Emissivity Coefficients Materials. https://www.engineeringtoolbox.com/emissivity-coefficients-d_447.html. Accessed 15 June 2018
20. Radiation Heat Transfer. https://www.engineeringtoolbox.com/radiation-heat-transfer-d_431.html. Accessed 15 June 2018
21. Fouladi, F., Henshaw, P., Ting, D.S.-K.: Enhancing smart grid realization with accurate prediction of photovoltaic performance based on weather forecast. Int. J. Environ. Stud. **70** (5), 754–764 (2013)
22. Ting, D.S.-K.: Basics of Engineering Turbulence. Academic Press, New York (2016)
23. Taylor, G.I.: The spectrum of turbulence. Proc. R. Soc. Lond. **164**(919), 476–490 (1938)
24. Figliola, R.S., Beasley, D.E.: Theory and Design for Mechanical Measurements, 4th edn. Wiley, New York (2006)
25. Yavuzkurt, S.: A guide to uncertainty analysis of hot-wire data. J. Fluids Eng. **106**, 181–186 (1984)
26. Tyagi, H., Liu, R., Ting, D.S.-K., Johnston, C.R.: Measurement of wake properties of a sphere in freestream turbulence. Exp. Thermal Fluid Sci. **30**, 587–604 (2006)
27. Dutta, S., Panigrahi, P.K., Muralidhar, K.: Experimental investigation of flow past a square cylinder at an angle of incidence. J. Eng. Mech. **134**(9), 788–803 (2008)
28. Dayem, A.M.A., Bayomi, N.N.: Experimental and numerical flow visualization of a single square cylinder. Int. J. Comput. Methods Eng. Sci. Mech. **7**, 113–127 (2006)
29. Okajima, A.: Strouhal numbers of rectangular cylinders. J. Fluid Mech. **123**, 379–398 (1982)
30. Hwang, R.R., Yao, C.-C.: A numerical study of vortex shedding from a square cylinder with ground effect. J. Fluids Eng. **119**(3), 512–518 (1997)
31. Bayraktar, S., Yayla, S., Oztekin, A., Ma, H.: Wall proximity effects on flow over cylinders with different cross sections. Can. J. Phys. **92**, 1141–1148 (2014)
32. Tariq, A., Panigrahi, P.K., Muralidhar, K.: Flow and heat transfer in the wake of a surface-mounted rib with a slit. Exp. Fluids **37**(5), 701–719 (2004)

`Predicting the Interior Conditions in a High Tunnel Greenhouse

Shreya Ghose and William Lubitz(✉)

School of Engineering, University of Guelph, 50 Stone Rd. E,
Guelph, ON N1G 2W1, Canada
wlubitz@uoguelph.ca

Abstract. Simple greenhouses constructed of polyethylene glazing over a steel frame with roll-up sides, called high tunnels, are used in horticulture to extend the growing season, improve crop quality, reduce disease, and reduce pest control issues. The high tunnel is effectively a passive solar structure installed over a crop that is grown using conventional, soil-based methods. Unlike large, hydroponic greenhouses, there are no fans or automated ventilation controls, nor any active heating or cooling systems. High-resolution measurements of air and soil temperatures, relative humidity, solar radiation, and wind speeds were recorded inside Quonset-style passive solar high tunnels, and adjacent open fields, for two growing seasons at two high tunnels in Guelph, Ontario. The side openings on one high tunnel were screened, while the other was left open. Lower than expected frost resistance was observed inside the high tunnels. A one-dimensional parametric energy model was developed to predict air and soil temperatures as a function of weather and high tunnel properties. The model was validated by comparing model predictions of air and soil temperatures to the measurements. A sensitivity study was conducted with the model using different combinations of parameter values to observe the effects of parameter choice on predicted high tunnel microclimate. Parametric models of high tunnel or greenhouse environments were found to be sensitive to choices of model parameters, such as soil, glazing, and thermal properties. The model and data collection was part of a longer study intended to benefit Canadian growers. An ability to accurately predict high tunnel microclimate will help growers choose crops best suited for high tunnels at their location. The data and models resulting from this study will also be useful for identifying methods of reducing energy costs in new high tunnel installations through structural changes and by modifying operating methods.

1 Introduction

High tunnels are simple frame greenhouses covered by a single layer of transparent polyethylene plastic sheeting. A high tunnel is heated by passive solar gain and ventilated by air movements forced by exterior winds and buoyancy effects. Ventilation is provided by rolling up the plastic sheeting from near ground level on each side of the structure. There is no active ventilation, heating or artificial lighting, making the high tunnel a relatively inexpensive type of greenhouse that can be readily deployed in locations without electrical or other utility infrastructure.

A. Vasel and D. S.-K. Ting (Eds.): EAS 2018, SPE, pp. 129–145, 2019.
https://doi.org/10.1007/978-3-030-00105-6_8

High tunnels are not a new technology. They have been used successfully for fruit and vegetable production around the world [1–3]. Different regions have different climate conditions and market opportunities, and because the interior climate in the high tunnel is entirely dependent on outside conditions, it can be difficult to transfer high tunnel experience in one location to another. Therefore, high tunnel performance, and feasibility for particular crops, is best tested and demonstrated at a regional level [2, 4].

A high tunnel can provide a higher temperature environment than would otherwise be available at a particular site. Air temperatures in high tunnels during sunny days in Nebraska, USA were shown to be above freezing even when ambient air temperatures stay below freezing [5]. The same study found that soil temperatures were significantly higher in the tunnel than outside [5]. These increases in air and soil temperature were also recorded in an unheated greenhouse in Guelph, Ontario during the winter of 2014 [6]. The effect can be substantial: air temperature differences between the high tunnel interior and outside of 30 °C or more occurred regularly during a winter measurement campaign during which outside temperatures repeatedly reached −18 °C.

Although Canadian growers have started to adopt high tunnels for high-value crops including strawberries and raspberries, there have been few regionally specific Canadian research and demonstration projects. There are many questions about the microclimate within high tunnels that are not answered in the literature, particularly in the context of the temperate, continental climate of southern Ontario, Canada. Particularly notable was an absence of any information about wind-driven ventilation, the efficacy of insect screening ventilation openings, and temperature stratification in high tunnels, which unlike most other greenhouse technologies, do not include roof-level ventilation [7]. This study attempted to fill this need by studying conditions in a set of research high tunnels at the Guelph Centre for Urban Organic Farming (GCUOF) on the University of Guelph campus (Fig. 1).

Fig. 1. High tunnels at the GCUOF in June 2016. The sides are opened for ventilation in the left-most tunnel, while they are closed in the center tunnel.

While there has been some study of wind-driven ventilation in high tunnel-like greenhouses [8] or single-bay greenhouses with side openings only, most of these studies have been simplified or reported limited results that could not be feasibly extrapolated to other high tunnel-type greenhouses. The impact of covering the

sidewall ventilation openings with insect screening, which would be expected to reduce pest pressures within the high tunnel reduces the air flow through the openings. There are no known studies that attempt to link the impact of insect screens on microclimate and insect control efficacy. Studies of the impact of insect screens on greenhouse ventilation usually focus on roof vents used in greenhouses, rather than side vent configurations typical of high tunnels [9]. Temperature stratification, particularly in low wind conditions and with climbing crops, is of particular concern yet is unstudied in this context. Summer solar radiation in Ontario can be higher than most plants need and generate excess heat in greenhouse and high tunnel environments [10]. High tunnel glazing blocks some incoming solar radiation but also reduces re-radiation in the infrared, causing overall elevated interior temperatures [11].

Organic production of fruits and vegetables is increasing in response to market demands driven by environmental and health concerns. Pest pressure, heat, and drought stresses in summer months are the most limiting factors on field organic vegetable production. The benefits of using high tunnels are not always clear: some recent research [4, 12] showed that occurrence of insects and disease in high tunnel production can be higher than in open fields. Other studies have shown that insect screening and shading materials can potentially improve growing conditions and increase crop yields in high tunnels, and that insect pests and pathogens can be reduced by using insect screening in greenhouse and high tunnel production systems [13–15]. Different insect screens have varying influences on environmental conditions, including light, temperature, humidity, and interior air movement ventilation rate [16–19]. A need was identified to directly and simultaneously measure the impact of insect netting on high tunnel interior microclimate and on insect populations, and provide guidance to growers planning to use high tunnels for various crops.

An important practical need for growers is the ability to predict the site-specific interior environment that will occur within a high tunnel that will result due to any particular outside weather conditions. This would allow growers to determine if it would seem feasible to use a high tunnel for a particular crop, before investing in constructing high tunnels. An important goal of this study was to develop a predictive model specific to high tunnels that could be used to predict the interior climate within a high tunnel for any particular location and weather conditions. Additionally, the process of development of the model may give insights into which heat transfer pathways are most significant, and this knowledge could be used to improve the design or operation of future high tunnels.

Greenhouse modeling studies are common in the literature and typically take one of three approaches. The first approach is to develop a series of equations that describe the heat exchange pathways within the greenhouse, discretize the equations, and then use this set of equations with time series weather data to predict parameters of interest such as greenhouse air or soil temperature as a function of time [20–22]. More recently, computational fluid dynamics (CFD) studies have been conducted of specific greenhouse configurations and conditions [23, 24]. A CFD simulation can provide a detailed understanding of a specific operating case, but challenges remain with using this approach to predict the effects of long-term weather patterns on the interior greenhouse climate. The practical usefulness of most CFD studies is limited by the inability to simulate long periods of weather and resulting greenhouse climate, and by difficulty of

obtaining the very high-quality measured data necessary to validate CFD results [25]. The third approach is to use so-called black box methods, including regression methods, artificial neural networks, or emerging machine learning techniques, to derive predictive relationships from long-term measured data [26]. These results are often limited, however, because the resulting predictive models may not be accurate if applied to even slightly different cases. Black box models are not generally able to provide much guidance about uncertainty levels in new cases, nor insight into how the parameters of interest are impacted by the forcing parameters [26].

The focus of this study was development of a lumped-capacitance, physics-based model of the high tunnel interior because of our practical need to be able to provide season-long site-specific predictions of high tunnel interior environments, and a desire to also gain insight into the effect of individual parameters (such as incoming solar radiation or glazing properties) on high tunnel interior conditions.

One of the most developed lumped-capacitance greenhouse models is the Gembloux Dynamic Greenhouse Climate Model (GDGCM), a multi-component semi-one-dimensional dynamic greenhouse climate model which calculates eight heat balances for the following greenhouse layers: cover, air, vegetation, soil surface, and four soil layers. It also includes a mass balance for the simulation of the relative humidity of the greenhouse air [20]. Although this model was developed for a soilless culture of tomato plants in large multi-span and actively ventilated European glasshouses, it was modified and adapted to be used for a naturally ventilated plastic greenhouse with a rose crop [20]. The approach used in this study is similar to that used to develop the GDGCM, namely, using a lumped-capacitance equation-based approach; however, this study does not specifically use the GDGCM model.

The work reported here is from a multiyear project that seeks to measure, model, and understand the microclimate that can be achieved within high tunnels in southern Ontario used for organic production of horticulture crops including tomatoes and bitter melon. Overall project goals include providing data, guidelines, and tools to growers that can be used to predict high tunnel microclimate, information needed to plant crops, and crop schedules. The project included high-resolution microclimate measurements within a set of research high tunnels located in Guelph, Ontario, Canada. The results presented here represent the first stage of development of a physics-based high tunnel greenhouse climate model that will eventually include more sophisticated representation of effects due to ventilation, moisture, and plants.

2 GCUOF High Tunnels

This study uses data from a set of research high tunnels located at the Guelph Centre for Urban Organic Farming (GCUOF; http://gcuof.wordpress.com/) in Guelph, Ontario, Canada (Fig. 1). Microclimate measurements were conducted in two high tunnels: one with unscreened side openings, and one with screened door and side openings. Screening was Econet 20/10 anti-insect mesh (Gintec Shade Technologies Inc., Windham Centre, ON, Canada), having a regular, rectangular mesh pattern, with a measured density of 40 threads per inch by 30 threads per inch. Optical porosity was measured at 40%. The site contained a total of six freestanding Quonset-style, steel-

framed high tunnels 7.2 m wide, 10.8 m long, and 3.8 m high at the roof peak (DeCloet Greenhouse Manufacturing, Delhi, ON, Canada) and three field plots in a 3 × 3 grid (Fig. 2). The high tunnels are glazed with a single layer of greenhouse-grade polyethylene sheeting, which can be manually rolled upward on the two long sides (but not the ends) to form ventilation openings that are the length of the high tunnel and 0.7 m tall. These openings were kept closed during low-temperature periods (typically when daytime maximum temperature was forecast to be less than 10 °C) and opened to prevent overheating during times of high solar radiation or ambient air temperature.

Crops of pea shoots, tomatoes, bitter melons, and edible chrysanthemum were grown in each plot during the growing season. The bitter melons [27] and tomatoes were climbing crops and trellised within the high tunnels. Three masts were placed in the center of the screened tunnel, central field, and unscreened tunnel (Fig. 2). These masts supported the following sensors:

- Air temperature sensors at heights of 0.3 m, 1.5 m at all sites, and also 3.0 m above ground in high tunnels only. Thermocouples were used in 2016 (Type E, in custom-made periodically ventilated housings), and thermistors (PT9 Pace Scientific series) were used in 2015.
- Soil temperature sensors at depths of 5 and 15 cm below ground. Thermistors in 2015 (Pace Scientific PT916), encapsulated Type E thermocouples in 2016.
- Solar radiation: Pace Scientific SRS-100 pyranometers.
- Airspeed and sonic air temperature: RM Young 81000 three-dimensional sonic anemometer with center of sensing volume positioned 1.00 m above ground.

Data was logged at one-minute intervals for all sensors using a combination of Pace XR5-SE and Campbell Scientific CR1000 data loggers. Data were recorded at the site from May 25, 2015 through September 20, 2015, and from May 14, 2016 through October 30, 2016. Additional measurements related to the crops trials, including times of side opening and closing, insect traps, plant growth, and crop yield, were also recorded throughout the growing seasons [27].

Table 1 shows the overall climate statistics as measured at the during the two field seasons in the open field, within the unscreened high tunnel, and within the screened high tunnel. Air and soil temperatures were higher within the high tunnels than in the open field. The mean temperatures within the screened tunnel were higher than in the unscreened tunnel, while the mean wind speeds observed in the screened tunnel were lower than in the unscreened tunnel. Air exchange is a significant heat exchange pathway when high tunnel sides are open, and the presence of insect screening over the side openings significantly reduced wind-driven air flow relative to an unscreened opening.

Figure 3 illustrates that the soil temperatures within the high tunnels were consistently warmer than the outside soil, particularly at the beginning and end of the growing seasons. Air temperatures showed a similar overall trend, but with much greater range and hourly/diurnal variability (Fig. 4). Typical conditions within the high tunnels included increased interior air temperatures during sunny periods, but air

Fig. 2. Aerial view of GCUOF high tunnel site. True north is upward. Monitoring stations were located as shown in a screened tunnel ("S"), an open field ("F"), and an unscreened tunnel ("U").

Table 1. Overall measured values for each field season.

Parameter	2015			2016		
	Field	Unscreened	Screened	Field	Unscreened	Screened
Mean air temperature (1.0 m) °C	16.2	17.3	18.3	18.4	19.8	21.0
Mean soil temperature (5 cm) °C	19.1	20.5	21.6	18.4	21.2	21.3
Mean soil temperature (15 cm) °C	18.8	19.9	20.9	18.0	20.2	20.8
Mean wind speed (1.0 m) m/s	0.26	0.06	0.04	0.19	0.10	0.08

temperatures returning to close to outdoor temperatures at night. This relatively low apparent thermal mass was consistently observed. It was also found that the high tunnels provided little frost protection without the addition of additional, non-transparent plant coverings. The effect of opening the high tunnel sides is also visible in Fig. 4, when interior air temperatures step downward at time of days 157.4, 158.4, and 159.5. The closing of the tunnel sides in the evening is also visible as abrupt, but smaller, increases in interior air temperature relative to outside temperature during the evenings.

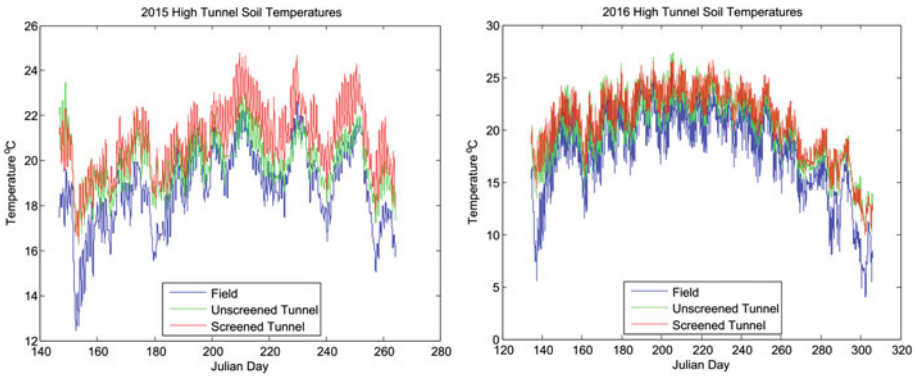


Fig. 3. Soil temperatures at 15 cm depth during the 2015 (left) and 2016 (right) measurement campaigns.

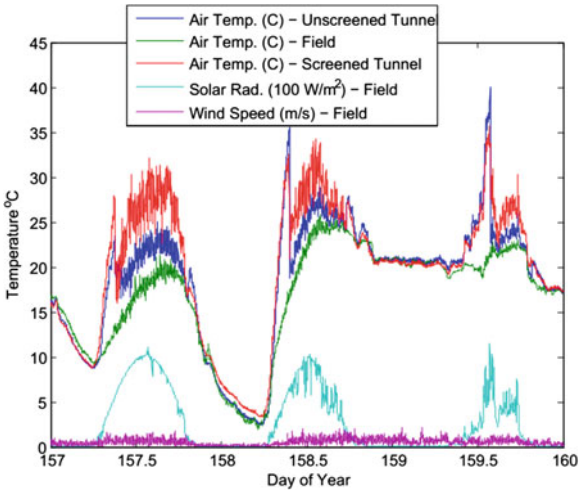


Fig. 4. Example conditions within high tunnels for 3 days in June 2015.

3 Model

The model of the GCUOF high tunnels was developed as a one-dimensional, lumped-capacitance thermal model, with heat transfer pathways as shown in Fig. 5. The greenhouse is modeled as a glazing layer, an airspace, and then several soil layers. Multiple soil layers were used to allow variation of soil temperature with depth at any given time, and past experience showed that this was necessary to achieve realistic soil conditions. Figure 5 also includes definitions of variables for material, thermal, and optical properties of each layer, and the values used for constants for these variables. The model described here is a simplified model and does not include a vegetation layer

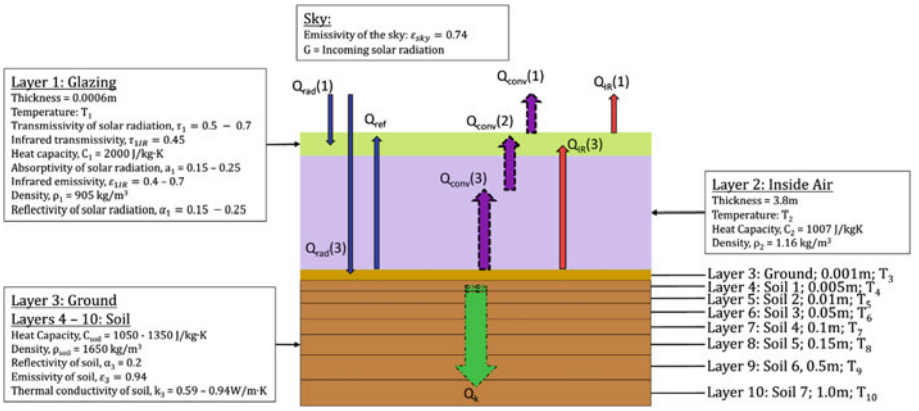


Fig. 5. Schematic representation of model zones, variables, and heat transfer pathways. Subscripts "k", "rad", "conv", and "IR" represent heat transfer by conduction, short-wave visible radiation, convection, and long-wave infrared radiation, respectively.

or energy transfer related to the vaporization or condensation of moisture. These limitations are discussed later.

The model was implemented by defining equations for each heat transfer pathway in Fig. 5. Assuming a one-minute time step during which all processes are quasi-static, the model was forced using measured outdoor meteorological data, which was recorded at the site using a one-minute averaging period. In the discussion that follows, the model is developed and used to predict conditions inside the high tunnels for periods when the sides of the tunnels remained closed for multiple days. Extension of the model for all operating periods would require additional development of air exchange relations that incorporate both air leakage during the closed condition (included in the following model) and bulk air exchanges when the sides are open (which is not yet included in the model described here).

The model includes radiation, conduction, and convection heat transfer between different high tunnel components. Past experience has suggested that all three forms of heat transfer must be included for a model to provide reasonable predictions of greenhouse interior conditions. The general form of the radiation equation for "gray bodies" is [28]

$$Q = \sigma A\left(T_1^4 - T_2^4\right) \tag{1}$$

The general form of natural or free convection equation [28] is

$$Q = hA(T_s - T_a) \tag{2}$$

and the general form of the conduction equation using Fourier's Law [28] is

$$Q = kA\frac{(T_2 - T_1)}{s} \tag{3}$$

where k is the thermal conductivity of the soil (W/mK), h is the convective heat transfer coefficient (W/m^{2}K), ε is the emissivity of the gray body, $\sigma = 5.67 \times 10^{-8}$ W/m^{2}k^4 (the Stephen–Boltzmann constant), and A is the footprint area of the high tunnel. Variables T_s and T_a refer to the surface temperature of the convecting body and the ambient air temperature, respectively, while variables T_1 and T_2 refer to the temperature of the two different zones or regions.

Applying conservation of energy, the following factors were identified as important to include in the model:

- Energy gained from incoming solar radiation,
- Total amount of energy absorbed by internal surfaces,
- Loss of energy to the surroundings, and
- Thermal capacity of soil.

The model was structured to predict the temperature of each of the ten layers as a function of incoming solar radiation, ambient air temperature, and ambient wind speed. The model was specifically set up to allow predictions of variables at high resolution in the time domain.

The model was validated by comparing model predictions of air and soil temperatures to the site measurements. The specific high tunnel considered has a single layer of polyethylene glazing stretched over the steel high tunnel frame. For the initial test cases, the sides are closed, and it is assumed that the door of the high tunnel was closed for the duration of the data collection. Effects on microclimate due to evapotranspiration by crop, relative humidity, latent heat transfer, and sensible heat transfer were not considered in this model. The glazing is assumed to have no large holes or tears which might affect the microclimate inside these tunnels, and each layer is horizontally homogenous.

The set of equations used to derive the rate of heat transfer between the ten model layers and the sky is given in Table 2. Numeric subscripts correspond to layer numbers shown in figure. The inputs provided to the model were the incoming solar irradiance G in W/m^2, the ambient temperature T_{amb}, in °C, and the wind speed (WS) in m/s. The temperature was converted to Kelvin (K) within the model as required by the equations.

Models were implemented in MATLAB (The Mathworks Inc., USA). The model was forward differenced in time with a 4-s time step to reduce drift due to any nonlinearities. Initial experience showed that the glazing layer needed to be included. The glazing has convection and radiation heat transfer processes occurring on both inside and outside surfaces. Incoming solar irradiance and reflected radiation from the soil surface are absorbed, transmitted, and reflected by the glazing.

Table 2. Model equations.

Description of Heat Transfer Equations	Equation	
Solar irradiance absorbed by glazing, $Q_{rad}(1)$	$Q_{rad}(1) = G \times \mathrm{a}_1$	(4)
Transmitted solar irradiance absorbed by soil, $Q_{rad}(3)$	$Q_{rad}(3) = G \times \tau_1 \times (1 - a_3)$	(5)
Solar radiation reflected from soil absorbed by glazing, Q_{ref}	$Q_{ref} = G \times \tau_1 \times a_3 \times \mathrm{a}_1$	(6)
Convection from soil to inside air, $Q_{conv}(3)$	$Q_{conv}(3) = h_c(3)(T_3 - T_2)$	(7)
Convection from inside air to glazing, $Q_{conv}(2)$	$Q_{conv}(2) = h_c(2)(T_2 - T_1)$	(8)
Convection from glazing to outdoor air, $Q_{conv}(1)$	$Q_{conv}(1) = h_c(1)(T_1 - T_{amb})$	(9)
Infra-red radiation from glazing to outdoor air, $Q_{conv}(1)$	$Q_{IR}(1) = \sigma\varepsilon_1(T_1)^4 - \varepsilon_{sky}(T_{rs})^4)$	(10)
Infra-red radiation from ground absorbed by glazing, $Q_{IR}(3)$	$Q_{IR}(3) = \sigma((1 - \tau_{1IR})(\varepsilon_3(T_3)^4 - \varepsilon_1(T_1)^4)$	(11)
Conduction heat transfer for each soil layer (n is the layer number), Q_k	$Q_k = \frac{k_3(T_n - T_{n-1})}{0.5dx_n + 0.5dx_{n-1}}$	(12)

The change in temperature of layer i over a single time step dt is then

$$dT_i = dt\frac{\sum_j Q_j}{c_i\rho_i dx_i} \tag{13}$$

where i is the layer number, j contains all the heat transfer paths into a specific layer, and c_i, ρ_i, and dx_i are, respectively, the heat capacity, density, and thickness of the layer.

Significant night energy losses can occur from greenhouses due to radiation exchange between the greenhouse and the night sky. Particularly on clear nights, the radiant sky temperature (the effective temperature of the sky for radiative heat transfer purposes) is often lower than ambient air temperature, and under the right conditions the resulting radiative heat transfer from the greenhouse to the sky can reduce the interior greenhouse temperature to one to two degrees below the outside air temperature [29]. Wang and Boulard [30] report radiant sky temperature T_{rs} can be modeled as

$$T_{rs} = F_{cn}T_{amb} + 0.0552(1 - F_{cn})T_{amb}^{1.5} \tag{14}$$

where T_{amb} is the ambient outside air temperature and F_{cn} is the fractional cloud cover, ranging from zero for a clear, cloudless sky to 1.0 for a fully overcast or cloud-covered sky.

Air leakage rates in closed greenhouses range from 0.5 to 5 air changes per hour [31]. In practice, some component of air exchange will be driven by ambient wind speeds. An air change correction was applied at each time step, in which a portion of interior air was replaced with air at outside air temperature. The rate of air changes (AC) per hour was approximated as 1 AC/h plus measured wind speed in m/s. (This relation has not been extensively tested.)

Table 3 lists the baseline values of several additional parameters required in the model. These were selected based on experience reported in prior literature. Soil

density, heat capacity, and conductivity were directly measured in the laboratory, and observed glazing transmissivity was based on a combination of prior literature values and differences in radiation measured inside and outside the high tunnels.

Table 3. Baseline parameter values for GCUOF high tunnel simulation model

Parameter name	Baseline values
Cloud-covered fraction of the sky, F_{cn}	0.5; partially cloudy skies
Transmissivity of glazing, τ_1	0.60 (Absorptivity = 0.20, reflectivity = 0.20)
Emissivity of glazing, ε_1	0.7
Thermal conductivity of soil, k	0.71 W/m K
Specific heat capacity of soil, c	1200 J/kg K
Convection from soil to inside air, $Q_{conv}(3)$	$h_c(3) = 3.4(\Delta T)^{0.33}$
Convection from inside air to glazing, $Q_{conv}(2)$	$h_c(2) = 2.21(T_2 - T_1)^{0.33}$ $(0.3 < T_2 - T_1 \leq 13.8\,^{\circ}\mathrm{C})$
Convection from glazing to outdoor air, $Q_{conv}(1)$	$h_c(1) = 7.2 + 3.84u$ (where u is wind speed in m/s)

4 Baseline Results

Three segments of data recorded minute-by-minute during the 2015 and 2016 growing seasons were used to test the performance of the model described in Sect. 3. For the 2015 growing season, a 91-hour segment of data from May 28 to June 1, 2015 was used. For the 2016 growing season, two segments of data were used in the simulations: a 54-hour segment from September 27 to September 29, 2016 (called "2016_1"), and a 65-hour segment from September 30 to October 3, 2016 (called "2016_2"). The cases are each several days long, to allow observation of the ability of the model to reproduce diurnal cycles, and the tunnel sides remained closed throughout each test period, to simplify the air exchange modeling. The initial air temperature was set equal to the measured air temperature at the start of the simulation, and initial temperatures for the 10 soil layers were determined by interpolating the measured soil temperatures at 5 cm and 15 cm at the simulation start time.

Figure 6 shows the predicted air and 5 cm (2 in.) soil temperatures for each of the three cases. It is notable that some time lagged responses are apparent in all three test cases. Additionally, while both air and soil temperatures often show a similar response (i.e., rising or falling at appropriate times) to the observed data, the magnitude of the values is often very different. In particular, it appears that the amount of heat flow into or out of the air or soil layer often remains incorrect. There are several possible explanations for the inconsistency between the predictions and the measured data. First, the model does not include thermal energy exchanges associated with the vaporization or condensation of water, which occurs within high tunnels. Prior researchers have shown that it is necessary to include these effects in the energy budget [20]. During sunny periods, plants transpire and moisture within the soil surface layer is evaporated. The addition of irrigation water would also be reasonably expected to impact soil and

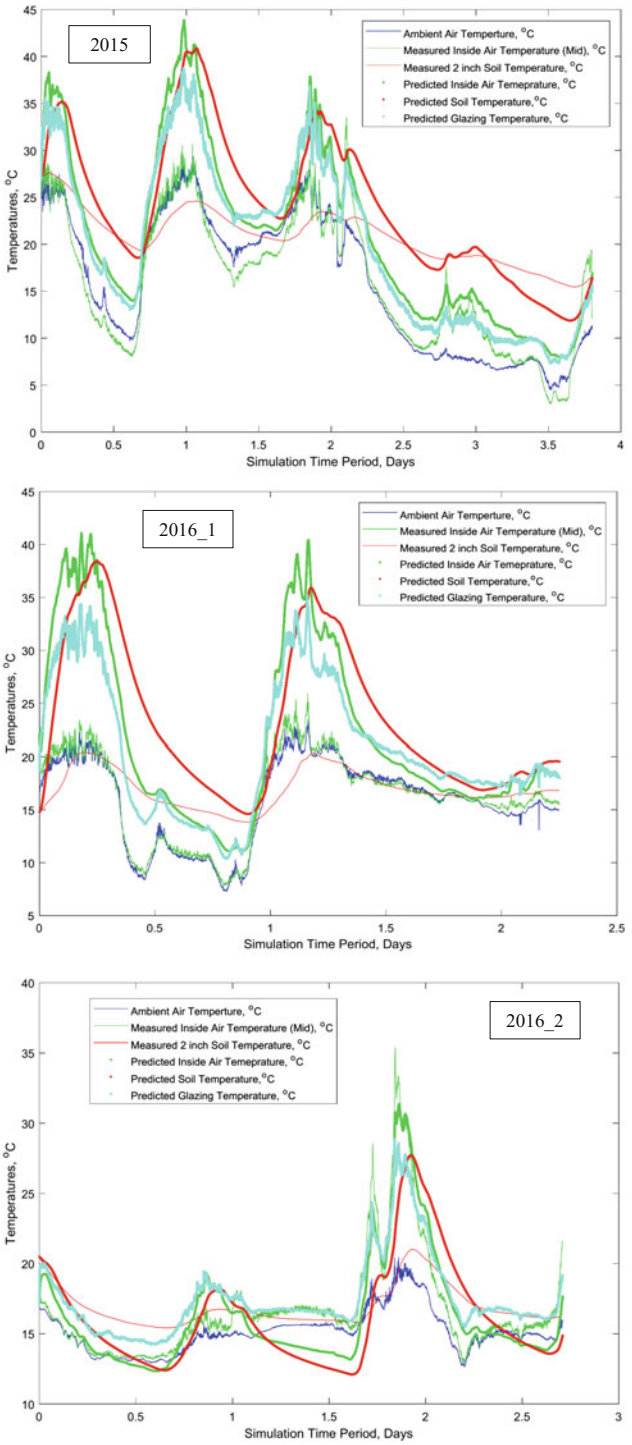


Fig. 6. Measured and predicted air and soil temperatures for baseline cases "2015" (top), "2016_1" (middle), and "2016_2" (bottom).

high tunnel interior temperatures. (Irrigation records were not available for this study.) Also, the varying presence and configuration of a plant canopy over a growing season should be included: plants shade the soil and absorb or give off heat, in a series of complex and difficult-to-model (or measure) processes.

The accuracy of the simulations was quantified by calculating the mean error (ME), mean absolute error (MAE), and root-mean-squared error (RMSE) over the entire simulation period for each case:

$$ME = \frac{\sum_{i=1}^{n} O_i - P_i}{n} \tag{15}$$

$$MAE = \frac{\sum_{i=1}^{n} O_i - P_i}{n} \tag{16}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (O_i - P_i)^2}{n}} \tag{17}$$

where n is the number of values, and O_i and P_i are observed and predicted values, respectively. Table 4 shows the calculated error statistics for predictions of the air and 5-cm soil temperature for each of the three cases.

Table 4. Error statistics for baseline simulations

Case	Air temperature °C			Soil temperature °C		
	MAE	ME	RMSE	MAE	ME	RMSE
2015	5.2	5.1	6.1	4.4	3.3	6.0
2016_1	5.7	5.7	7.8	6.5	6.5	8.4
2016_2	1.4	−0.4	1.7	2.1	−0.7	2.5

5 Sensitivity Study

The model outlined in Sect. 3, using the baseline parameter values in Sect. 4, was shown to predict some features of the high tunnel interior conditions, but there is still significant room for improvement. While containing dozens of constants and variables, the model is still very simple and does not yet include some likely important phenomena. In addition, there are already nonlinear relationships (e.g., radiative heat transfer) within the model, and the coupling between many variables is complex. Under these circumstances, it is difficult to estimate *a priori* whether an inaccuracy in a particular parameter (such as a material property) would have a negligible, or alternatively, very large impact on the simulation results. Since most of the parameters in a greenhouse climate model are estimated, and may have a wide range of values in different, real high tunnels, this is an important question. The overall question, which

unfortunately will not be able to be answered here, is what prediction accuracy that can be achieved by equation-based lumped-capacitance models when trying to predict the conditions inside real high tunnels? The authors found essentially no discussion of this important topic in the literature. (It is also worth noting that this question can, and should, also be asked of CFD or black box simulations, since these are also similarly based on estimated parameters and other necessary simplifications.)

To begin to answer some of these questions, a series of sensitivity studies were completed. The parameters in Table 5 were selected due to either a relative lack of information in the literature, a lack of agreement on appropriate values or discussion of details, or because they would be expected to be particularly important for the response of the model over time.

Effects of varying the cloud-covered fraction of the sky, transmissivity and emissivity of the glazing, and thermal conductivity (with correlating density and heat capacity) of the soil were monitored by changing the values of a single parameter one at a time, while holding all other parameters constant at the baseline. The variations of these parameters are shown in Table 5 and were applied to each of the three test cases. As values of transmissivity were varied, corresponding absorptivity and reflectivity values were scaled proportionally relative to the values used in the baseline simulation. Only one variable was changed at a time, and all other variables were kept at baseline values.

Inspection of Table 5 shows that the model error statistics were sensitive to all of the variables considered. Changing glazing emissivity and transmissivity produced larger ranges of error values than changing the sky cloud cover or soil conductivity. It is also notable that changing a variable did not always cause the same type of change in results for the three test cases.

Table 5 also shows error statistics for three specific runs of the model: the baseline case, a clear sky case where $F_{cn} = 0$, and a low glazing transmissivity case with $\tau_1 = 0.25$. Both of the latter cases had overall lower MAE and ME air temperature values than the baseline case, although large improvements in ME and MAE for the 2015 and 2016_1 cases were offset by small increases for the 2016_2 case. (Additionally, they also both had much lower error levels for soil temperature predictions compared to the baseline case.) It is worth noting that the model predictions of the 2016_1 and 2016_2 cases are qualitatively very different (Fig. 4) even though these two cases represent two periods separated by only one day.

While some combinations of parameters produced more accurate simulations than others, these data (and other simulations trials not shown here) strongly suggest that there is no one "correct" set of parameters for a specific high tunnel. Achieving accurate predictions of one set of test data cannot be assumed to mean that the model will accurately predict conditions at a different time of year, or with a different crop (or as shown in this case, even a few days later). This is an important point, since most examples of model validation found in the literature (whether lumped-capacitance, CFD or black box) used only one or a few days of data.

While development is continuing on the model to include moisture and plant canopy effects, it is in the context that perfect time series simulation of the high tunnel

Table 5. Sensitivity study parameters with ranges of values, and resulting ranges of ME and MAE values for air temperature (°C) within the high tunnel.

Varying parameter name	Range of values	2015 MAE	2016_1 MAE	2016_2 MAE	2015 ME	2016_1 ME	2016_2 ME
Cloud-covered sky fraction, F_{cn}	0, 0.25, 0.5, 0.75, 1	4.2–6.3	4.8–7.0	1.3–2.0	4.0–6.2	4.6–7.0	−1.7–1.0
Transmissivity of glazing, τ_1	0%, 25%, 50%, 75%, 100%	2.2–9.1	1.8–11.3	1.4–3.9	−1.5–9.1	−3.4–11.3	−3.8–1.7
Emissivity of glazing, ε_1	0, 0.25, 0.5, 0.75, 1	3.6–14.0	4.0–16.6	1.4–10.7	3.4–13.9	3.5–16.6	−2.5–10.7
Thermal conductivity of soil, k	0.099, 0.504, 0.587, 0.712, 0.934 W/m K	3.6–5.2	3.4–8.3	1.4–4.5	2.9–5.1	3.0–5.7	−2.6–−0.4
Case name							
Baseline	$F_{cn} = 0.5$ $\tau_1 = 0.6$ $\varepsilon_1 = 0.7$ $k = 0.71$ W/m K	5.2	5.7	1.4	5.1	5.7	−0.4
Clear sky	$F_{cn} = 0$ $\tau_1 = 0.6$ $\varepsilon_1 = 0.7$ $k = 0.71$ W/m K	4.2	4.8	2.0	4.0	4.6	−1.7
Low transmissivity	$F_{cn} = 0.5$ $\tau_1 = 0.25$ $\varepsilon_1 = 0.7$ $k = 0.71$ W/m K	2.2	1.8	2.5	1.2	0.5	−2.3

environment will not be achievable, and that the effort should focus on accurately predicting practical values of most interest to growers, such as overall season length, maximum or minimum temperatures, seasonal degree days, and early season soil temperatures.

6 Conclusions

Interior and outside conditions were recorded in two high tunnel greenhouses in Guelph, Ontario. It was shown that the tunnels had less thermal mass than expected and provided little additional frost protection, but did successfully produce higher air and soil temperatures compared to outside on the order of a few degrees Celsius. Covering side ventilation openings with insect screening resulted in a further increase in average air and soil temperatures. It was found that the variation of the interior climate within the high tunnels could be modeled using a lumped-capacitance model; however, significant differences between predicted and observed conditions remained. Sensitivity studies using the model found that varying individual model parameters across a range of practical possible values could produce large changes in overall prediction accuracy. The results presented here strongly suggest the need to include moisture, plant canopy, and additional ventilation effects in the next generation of the model, and also suggest that there will be fundamental limits to the accuracy that can be achieved by modeling such complex, multi-variable systems using models that cannot practically reproduce the true variability, complexity, and full range of processes that occur in the real system.

Acknowledgements. This project was financially supported by Agriculture and Argi-food Canada through the Agricultural Adaptation Council and by the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) University of Guelph Partnership. The work reported here would not have been possible without the efforts of all those who contributed to the larger high

tunnel project, including Youbin Zheng, Dave Llewellyn, Yun Kong, Evan Elford, Mary Ruth McDonald, Ralph Martin, Martha Gay Scroggins, Rene van Acker, and all the summer student volunteers who supported crop cultivation and research at the GCUOF.

References

1. Carey, E.E., Jett, L., Lamont Jr., W.J., Nennici, T.T., Orzolek, M.D., Williams, K.A.: Horticultural crop production in high tunnels in the United States: a snapshot. HortTechnology **19**, 1 (2009)
2. O'Connell, S., Rivard, C., Peet, M.M., Chris, C., Louws, F.: High tunnel and field production of organic heirloom tomatoes: yield, fruit quality, disease, and microclimate. HortScience **47**, 1283–1290 (2012)
3. Wittwer, S.H., Castilla, N.: Protected cultivation of horticultural crops worldwide. HortTechnology **5**, 6–23 (1995)
4. Wallace, R.W., et al.: Lettuce yield and quality when grown in high tunnel and open-field production systems under three diverse climates. HortTechnology **22**, 659–668 (2012)
5. Wien, H.C.: Microenvironmental Variations within the high tunnel. HortScience **44**, 235–238 (2009)
6. Lubitz, D.: Reducing Canadian greenhouse energy costs using highly insulating glazing. In: World Renewable Energy Conference, London, UK, 3–8 August 2014
7. Boulard, T.: Recent trends in protected cultivations – microclimate studies: a review. Acta Hort. **057**, 15–28 (2012)
8. Zhao, X., Carey, E.E.: Summer production of lettuce, and microclimate in high tunnel and open field plots in Kansas. HortTechnology **19**, 113–119 (2009)
9. Katsoulas, N., et al.: Effect of vent openings and insect screens on greenhouse ventilation. Biosyst. Eng. **93**(4), 427–436 (2006)
10. Zheng. Y., Llewellyn, D., Dixon, M.: Investigation of the feasibility to improve the lighting environment under hanging baskets in greenhouse flower production. Technical report for Flowers Canada (2013)
11. Papadakis, G., et al.: Radiometric and thermal properties of, and testing methods for, greenhouse covering materials. J. Agric. Eng. Res. **2000**(77), 7–38 (2000)
12. Wien, H.C., Pritts, M.P.: Use of high tunnels in the Northeastern USA: adaptation to cold climates. Acta Hort. **807**, 55–59 (2009)
13. Ben-Yakir, D., et al.: Colored shading nets impede insect invasion and decrease the incidences of insect-transmitted viral diseases in vegetable crops. Entomol. Exp. Appl. **144**, 249–257 (2012)
14. Giordano, et al.: Effects of plastic screens on virus infection, yield and qualitative characteristics of small tomatoes. Acta Hort. **614**, 735–740 (2003)
15. Legarrea, S., et al.: Comparison of UV-absorbing nets in pepper crops: spectral properties. Photochem. Photobiol. **86**, 324–330 (2010)
16. Boulard, T., et al.: Improving air transfer through insect proof screens. Acta Hort. **893**, 289–296 (2011)
17. Muñoz, P., et al.: Natural ventilation of multi-span tunnel greenhouses with and without insect-proof screens. Acta Hort. **559**, 263–269 (2001)
18. Molina-Aiz, et al.: Effects of insect-proof screens used in greenhouse on microclimate and fruit yield of tomato (Solanum lycopersicum L.) in a Mediterranean climate. Acta Hort. **927**, 707–714 (2012)

19. Teitel, M.: The effect of screens on the microclimate of greenhouses and screenhouses – a review. Acta Hort. **719**, 575–586 (2006)
20. Mashonjowa, E., Ronsse, F., Milford, J.R., Pieters, J.G.: Modelling the thermal performance of a naturally ventilated greenhouse in Zimbabwe using a dynamic greenhouse climate model. Sol. Energy **91**, 381–393 (2013)
21. Roy, J.C., Boulard, T., Kittas, C., Wang, S.: PA—precision agriculture: convective and ventilation transfers in greenhouses, part 1: the greenhouse considered as a perfectly stirred tank. Biosyst. Eng. **83**(1), 1–20 (2002)
22. Zhang, Y., Mahrer, Y., Margolin, M.: Predicting the microclimate inside a greenhouse: an application of a one-dimensional numerical model in an unheated greenhouse. Agric. Forest Meteorol. **86**(3), 291–297 (1997)
23. Tong, G., Christopher, D.M., Li, B.: Numerical modelling of temperature variations in a Chinese solar greenhouse. Comput. Electron. Agric. **68**(1), 129–139 (2009)
24. Nebbali, R., Roy, J.C., Boulard, T.: Dynamic simulation of the distributed radiative and convective climate within a cropped greenhouse. Renew. Energy **43**, 111–129 (2012)
25. Zhong, W., Yu, A., Zhou, G., Xie, J., Zhang, H.: CFD simulation of dense particulate reaction system: approaches, recent advances and applications. Chem. Eng. Sci. **140**, 16–43 (2016). https://doi.org/10.1016/j.ces.2015.09.035
26. Picheny, V., Trépos, R., Casadebaig, P.: Optimization of black-box models with uncertain climatic inputs-Application to sunflower ideotype design. PLoS ONE **12**(5), 1–15 (2017). https://doi.org/10.1371/journal.pone.0176815
27. Kong, Y., Llewellyn, D., Schiestel, K., Scroggins, M.G., Lubitz, D., McDonald, M.R., Van Acker, R., Martin, R.C., Zheng, Y., Elford, E.: High tunnels can promote growth, yield, and fruit quality of organic bitter melons (Momordica charantia) in regions with cool and short growing seasons. HortScience **52**(1), 65–71 (2017)
28. Cengel, Y.A., Boles, M.A.: Thermodynamics: An Engineering Approach, 7th edn. McGraw-Hill, New York (2011)
29. Castilla, N.: Greenhouse Technology and Management, 2nd edn. CABI International, Wallingford (2013)
30. Wang, S., Boulard, T.: Predicting the microclimate in a naturally ventilated plastic house in a Mediterranean climate. J. Agric. Eng. Res. **75**(1), 27–38 (2000)
31. USDA: Virtual Grower. Version 3 Manual. U.S. Department of Agriculture, Agricultural Research Service (2014). http://www.ars.usda.gov/services/software/download.htm?softwareid=309

Assessment of a Heat Pump-Based Wastewater Heat Recovery System for a Canadian University Campus

Colton Chow(✉) and Jean Duquette

Department of Mechanical and Aerospace Engineering, Carleton University, 1125 Colonel By Drive, Ottawa, ON, Canada
colton.chow@carleton.ca

Abstract. Domestic hot water use in Canada accounts for roughly 19% and 8% of total residential and commercial building energy use, respectively. After use by building occupants, domestic hot water is disposed of in the sewer system where it mixes with other wastewater streams, typically reaching temperatures as high as 25 °C. Thus, an opportunity exists for recovering low-grade heat from wastewater and using it for meeting local thermal loads. This study uses the Carleton University campus, located in Ottawa, Canada, as a case study for wastewater heat recovery. The campus wastewater energy resource is estimated using a novel method, and a feasibility analysis is conducted on an industrial-scale heat pump that recovers waste heat from the university's main sewer outflow pipe. The recovered heat is used for heating water at the university's main athletics facility. Results show that installing a wastewater heat recovery system decreases greenhouse gas emissions and total annual costs by approximately 95% and 8%, respectively, relative to a base case natural gas heating system. Low-grade wastewater heat is an underutilized energy source that is both sustainable and accessible in most urban centres across Canada. Increasing the capacity of wastewater heat recovery systems has the potential to significantly increase energy savings and decrease greenhouse gas emissions in our energy system.

Nomenclature

Abbreviations

GHG	Greenhouse gases
COP	Coefficient of performance
BC	Base case scenario
SHR	Sewer heat recovery scenario
BCRE	Best-case recoverable energy estimation
WCRE	Worst-case recoverable energy estimation

Variables

$\dot{V}(t)$	Building water consumption (m^3/hr)
$\dot{V}_{sewer}(t)$	Hourly variation of flow in the main sewer outflow pipe (m^3/hr)
V_{heated}	Volume of water heated by a building in a set time period (m^3)
Q	Energy used to heat hot water in buildings (kJ)

A. Vasel and D. S.-K. Ting (Eds.): EAS 2018, SPE, pp. 146–167, 2019.
https://doi.org/10.1007/978-3-030-00105-6_9

ρ	Density of water (kg/m^3)
C_p	Heat capacity of water (kJ/kgK)
T_{SP}	Temperature of the hot water tanks on campus (°C)
T_{mains}	Temperature of fresh water in the water mains (°C)
ϕ	Heating ratio
V_{total}	Total volume of water used in a building during a set time period (m^3)
T_{out}	Temperature of the outflow from a single building (°C)
$\dot{Q}_{Sewer}(t)$	Energy in the main sewer outflow pipe (kW)
$T_{sewer}(t)$	Hourly variation of temperature in the main sewer outflow pipe (°C)
$\dot{Q}_{recoverable}(t)$	Hourly variation of energy that is recoverable from the main sewer outflow pipe (kW)
$T_{discharge}$	Temperature at which the effluent leaves the Carleton campus and enters the city sewers (°C)
$\dot{Q}_{out,BC}(t)$	Energy output to meet the domestic hot water load in the base case scenario (kW)
$\dot{Q}_{fuel}(t)$	Energy content of the fuel entering the boiler in the base case scenario (kW)
η_{BC}	Average efficiency of the boiler in the base case scenario (%)
$\dot{W}_{in,ep}(t)$	Energy required to pump effluent from the wet well to the heat pump evaporator (kW)
$\dot{Q}_{evap}(t)$	Energy absorbed by the heat pump evaporator (kW)
$T_{LLT}(t)$	Temperature at which the effluent leaves the heat pump evaporator (°C)
g	Rate of acceleration due to gravity (m/s^2)
ΔH	Effluent pump system head loss (m)
η_{ep}	Effluent pump efficiency (%)
f	Friction factor
L	Effluent pump system pipe length (m)
d	Effluent pump system pipe diameter (m)
v	Fluid velocity in the pipe (m/s)
K	Minor loss coefficient
$\dot{V}_{\max,ep}$	Maximum volumetric flow rate through the effluent pump system (m^3/s)
v_{econ}	Economic velocity of water (m/s)
$\dot{Q}_{out,SHR}(t)$	Energy output to meet the domestic hot water load in the sewer heat recovery scenario (kW)
$\dot{W}_{in,HP}(t)$	Energy required to drive the heat pump compressor (kW)
$COP(t)$	Heat pump coefficient of performance (COP)
AC	Total annual cost ($/yr)
NPC	Net present cost ($)
i	Discount rate (%)
C_c	Capital cost ($)
M_c	Maintenance cost ($/yr)
F_c	Variable fuel cost ($/yr)
j	Number of periods in net present cost calculation
m_{CO_2e}	GHG emissions in equivalent metric tonnes of CO_2 (tonnes)

AEO	Annual energy output (MWh)
e_{CO_2e}	Greenhouse gas emission factor (kg/MWh)
k	Fuel escalation rate (%)
n	System lifetime (yr)

1 Introduction

Canada has recently set a greenhouse gas (GHG) emission reduction target of 30% below 2005 levels by the year 2030 [1]. Current projections by Environment and Climate Change Canada indicate that Canada will most likely miss this target, with current policies in place resulting in GHG emissions that are 42% above the stated target [2]. Thus, greater investments are needed in innovative energy technologies that have the capacity to significantly decrease GHG emissions below current levels [3]. Approximately 20% of all primary energy use in Canada occurs in buildings [4], representing roughly 12% of total GHG emissions [5]. Of this total building energy use, approximately 19% and 8% can be attributed to domestic hot water use in the residential and commercial sectors, respectively [6]. Domestic hot water is discharged to the sewer system with most of its original energy content. In the sewer system, it mixes with other wastewater streams producing wastewater temperatures that are typically as high as 25 °C [7]. This represents a renewable source of low-grade heat that can be recovered to meet local thermal loads.

Studies have been conducted to assess the availability of low-grade thermal energy in urban sewer systems. Cipolla and Maglionico [8] conducted an assessment of the wastewater thermal energy resource from sewers located in the city of Bologna, Italy. They recorded temperature and flow data on four branches of the sewer system, and at the wastewater treatment plant. The data showed that effluent temperatures varied between 10 and 14 °C in the winter and 18 and 22 °C in the summer. They also showed that although flow in the sewer is a function of population size and of the occupant's behaviour, the same cannot be said of the effluent temperature. The effluent temperature was found to be independent of the size of the population that supplies it. Bertrand et al. [9] conducted an urban-scale assessment of the potential energy savings from using in-building wastewater heat recovery systems, based on pinch analysis. Their results showed that heat recovery with a heat pump could yield electricity savings of up to 28% and 41% for single-family homes and multifamily buildings, respectively.

Studies have also been conducted to assess the performance of various wastewater heat recovery systems. Wallin and Claesson [10] designed an experiment to test a heat pump and vertical falling film inline heat exchanger for recovering low-temperature heat from drain water. Their analysis shows that a large thermal contact resistance exists in the heat exchanger, which inhibits heat transfer from the drain water. However, at the tested flow rates, the experimental setup was able to recover more than 25% of the available heat. Dong et al. [11] designed and tested a heat pump water heater that recovers energy from shower drain water. They found the system was able to reduce energy consumption by 70% compared to a conventional electric water heater.

Hepbasli et al. [7] conducted a review of wastewater source heat pump systems. They concluded that although cities on average send almost 40% of the total thermal energy they generate into the sewer system as low-grade waste heat, there are often a number of challenges associated with recovering this heat due to equipment fouling issues. While these issues may be resolved by using treated wastewater, they found the cost of purchasing and installing piping to and from treated wastewater locations to be prohibitive. Furthermore, of the studies they reviewed, they found that the coefficient of performance (COP) of heating ranged from 1.8–10.6 and the COP of cooling ranged from 2.2–5.4.

There are currently an estimated five hundred wastewater heat recovery systems in the world, many of which are located in Scandinavia [7]. Only a few of these systems are in operation in North America. One example is in the city of Vancouver, Canada, which adopted a wastewater heat recovery system with a heat pump as part of their 2010 Olympic Village development [12]. The system recovers heat from untreated urban wastewater with a heat pump and distributes it to 14 local buildings through a district heating grid [13]. The wastewater heat recovery system is able to meet 70% of the annual energy demand and has reduced GHG emissions by 50% [12]. Another North American wastewater heat recovery system was installed as part of a new addition to the Denver Museum of Nature & Science in 2013 [14]. Im et al. [14] conducted an assessment of this system and found that it reduced CO_2 emissions by 41%, and saved 34% of the energy costs relative to a baseline HVAC system consisting of a natural gas boiler and water cooled chiller. They also concluded that the economic feasibility of the system was highly sensitive to the length of piping that is required between the wastewater resource and the museum.

Most of the studies described above were conducted using measured historical temperature and flow rate data from a given wastewater stream. While using this data increases the accuracy of estimating the available resource at a given site, it requires the foresight to preemptively collect the data. In most cases, this is impractical as measurement and data acquisition equipment is needed. Furthermore, a minimum waiting time of 1 year is typically required to collect a representative data sample. No studies have been identified that use existing utility consumption data, such as metered water or natural gas, to assess the available wastewater heat resource at a specified site. Utility consumption data is typically available over long time horizons and can be obtained directly from local building operators and/or utilities.

The objectives of this research are to (1) present a novel methodology that utilizes local utility consumption data in buildings to estimate the potential for recovering heat from downstream sewer wastewater, and (2) assess the economic and environmental impacts of switching from a conventional natural gas water heater to a heat pump-based sewer heat recovery system for heating domestic hot water. The Carleton University campus, located in Ottawa, Ontario, is chosen as the case study for the analysis. Carleton University is a mid-size university with approximately 29 500 students enrolled, and a campus spanning roughly 62 hectares [15]. Further to estimating the available sewer waste heat resource on this campus, two energy models are constructed corresponding to a base case (BC) scenario and a sewer heat recovery (SHR) scenario. Both scenarios represent systems that are capable of meeting the domestic hot water load in the main athletics facility located on campus. The BC scenario describes a

baseline system that comprises a natural gas boiler plant, whereas the SHR scenario comprises an industrial-scale heat pump that recovers heat from the campus' main sewer outflow pipe. Both energy models are compared with respect to the total annual cost and GHG emissions.

2 Methodology

2.1 Waste Heat Resource Assessment

The Carleton University campus is comprised of 42 buildings that are connected to a centralised sewer system. A separate stormwater system is used to collect rainwater. All buildings at the university are metered for electricity, water, and steam consumption. Steam is supplied to the buildings via a district heating grid and is used primarily for space and domestic hot water heating. Both the water that is consumed in buildings and the condensate discharged from the district heating grid are returned to the sewer system. One exception to this occurs during hot summer days when cooling towers are in operation. Cooling towers consume a considerable amount of water, most of which is rejected to the local environment instead of the sewer. In this study, the water use of buildings with cooling towers is discarded on days when the cooling towers are operational. On this basis, a building's water consumption is equivalent to the building's sewer flow, as all water consumed in the building is discharged to the sewer.

To assess the recoverable sewer waste heat resource, an estimate of the hourly flow rate and temperature in the university's main sewer outflow pipe is needed. In this study, a novel method is proposed for estimating these parameters based solely on monthly (and hourly) metered utility data collected from the individual campus buildings. The steps taken to estimate the recoverable energy in the sewer are depicted in Fig. 1 and described in detail in the remainder of this section.

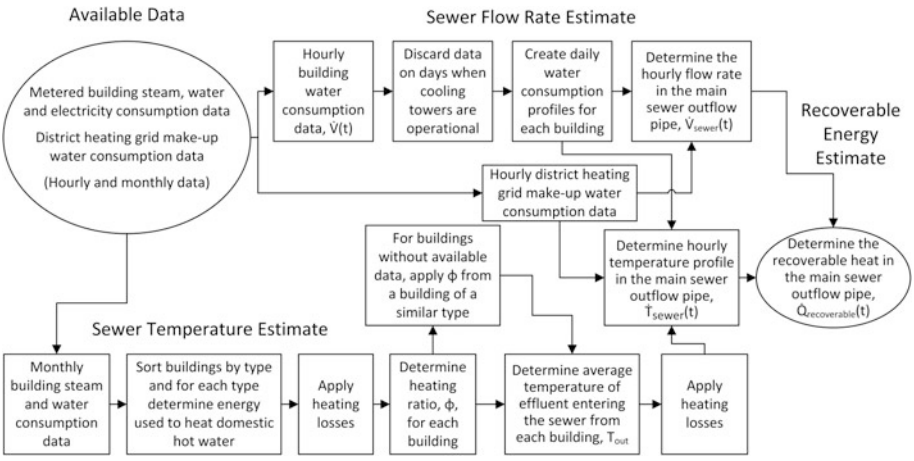


Fig. 1. Steps taken to estimate recoverable energy in the main campus sewer outflow pipe

The hourly flow rate in the main sewer outflow pipe is estimated by summing the individual daily water consumption profiles of each building on campus, as shown in Fig. 1. Four different daily profiles are used in this study as these are assumed to be representative of the four most common building occupancy conditions occurring at Carleton University. Daily profiles are used because the water and steam data required to make an estimate of the available energy is not available consistently throughout the year. Furthermore, using daily profiles reduces the computational burden of the model, compared to using hourly data from a year and extrapolating missing data points. Building occupancy is used to distinguish between water consumption profiles because it has the most significant effect on water use [16]. The daily water consumption profiles selected correspond to a weekday during the fall/winter academic term, a weekend during the fall/winter academic term, a weekday during an academic break, and a weekend during an academic break. At Carleton University the fall/winter academic term runs from September to April. Academic breaks occur throughout the year, such as from May to August for the summer break, or in the last week of October for the fall reading week. The building water consumption profiles are based on hourly building water consumption data, $\dot{V}(t)$, which was measured for a period of 1 year from November 2016 to December 2017. The building water consumption profiles show the building water use during each hour throughout the day. The building water consumption profiles are calculated by separating each building's data into categories based on the four building occupancy conditions described above. For each building occupancy condition, the average flow is calculated during each hour of the day. The average flow is shown in the building flow profile, for the corresponding hour, and combining all hours of the day makes up the building flow profile. Building water consumption includes both domestic water consumption and condensate that is returned to the sewer system from the district heating grid. In the latter case, the measured makeup water that replaces the discharged condensate is used for the evaluation. Daily water consumption profiles are determined for every building on campus except for four buildings that do not have data available. For these four buildings without data, a profile is created based on a similar building with respect to type and size. The hourly variation of flow in the main sewer outflow pipe for the four daily occupancy profiles mentioned above is shown in Fig. 2.

Also shown in Fig. 1 is the method used for estimating the hourly wastewater temperature in the main sewer outflow pipe. This method is based on the following two sources: 1) the district heating grid condensate that is returned to the sewer after exchanging heat in each building's substation, and 2) the metered energy used to heat domestic hot water in each building on campus. For the former, the condensate volume is assumed to be equivalent to the makeup water entering the district heating grid and the condensate temperature is assumed to be constant at 100 °C [17]. Each campus building comprises a substation from which steam from the district heating grid exchanges energy with the building energy distribution system. Energy from the district heating grid is exchanged via a steam-to-water heat exchanger that supplies both the building hydronic space heating loop, and the domestic hot water loop. Since building-based energy metres provide the total thermal energy flowing into any given building, energy metre data in this study is only considered over a 4-month period from

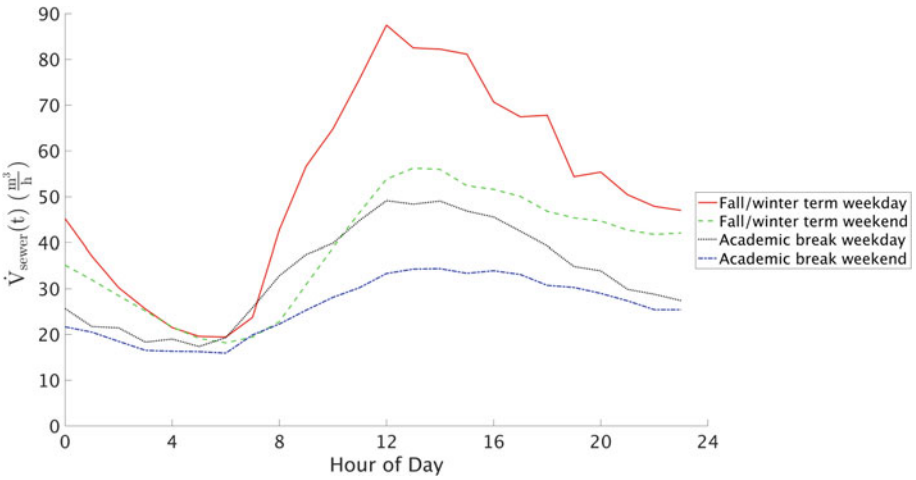


Fig. 2. Hourly variation of flow, $\dot{V}_{sewer}(t)$, in the main sewer outflow pipe over the course of an average day. Variations are shown for both weekdays and weekends during the winter academic term and the summer academic break. Hours 0–24 represent a 24-h period from 12 am to 12 am.

the beginning of June to the beginning of October, when no space heating loads are present [18]. Using data over this 4-month period ensures that the steam used in the buildings represents the energy used to heat domestic hot water, which is generally the only thermal load present. For buildings with available steam data, the following expression is used to calculate the volume of domestic water that is heated over the 4-month period from June to October:

$$V_{heated} = \frac{Q}{\rho C_p [T_{SP} - T_{mains}]} \tag{1}$$

where Q, ρ, and C_p represent the metered energy used to heat domestic hot water, the density of water, and the heat capacity of water, respectively. T_{SP} and T_{mains} represent the set point hot water tank temperature and the water mains temperature, respectively. In the current analysis, water is assumed to enter at 4 °C from the water mains and is heated to 55 °C in building-based hot water tanks. These values are consistent with those provided from the university's facilities management and planning group [18].

To determine the relative amount of incoming mains water that is heated in any given building for domestic hot water use, the following heating ratio is used:

$$\phi = \frac{V_{heated}}{V_{total}} \tag{2}$$

where V_{total} represents the total volume of measured incoming mains water over this same 4-month period.

For buildings that do not have available steam data, the heating ratio from a building of a similar type is used. A building type can be described as a building that provides a specific function (e.g., athletics facility, cafeteria, student residence, library, laboratory, etc.). Since most buildings on campus provide a variety of functions, no two buildings are exactly alike. A crude approach is thus used in the current study for matching building types, which consists of finding the best fit with respect to building function and floor space designated for each function. The temperature of the wastewater entering the sewer from each building is estimated using a first law energy balance approach and Eq. 3 as follows:

$$T_{out} = \phi T_{SP} + (1 - \phi) T_{mains} \tag{3}$$

The hourly available energy in the main sewer outflow pipe, $\dot{Q}_{sewer}(t)$, is estimated by summing the energy in each building's wastewater flow and is expressed as

$$\dot{Q}_{sewer}(t) = \sum_{i} \dot{V}(t)_i \rho C_p T_{out,i} \tag{4}$$

where $\dot{V}(t)$ represents the hourly measured domestic water flow rate in a given building on campus, and the subscript i represents an individual building. The hourly wastewater temperature in the main sewer outflow pipe, $T_{sewer}(t)$, can be calculated from Eq. 5 as follows:

$$T_{sewer}(t) = \frac{\dot{Q}_{sewer}(t)}{\dot{V}_{sewer}(t) \rho C_p} \tag{5}$$

where $\dot{V}_{sewer}(t)$ represents the hourly effluent flow rate in the main sewer outflow pipe as shown in Fig. 2. The hourly variation of wastewater temperature in the main sewer outflow pipe for the four daily occupancy profiles depicted in Fig. 2 is shown in Fig. 3.

Three types of heating losses are accounted for in this analysis, including hot water tank standby losses, point-of-use losses, and sewer system distribution losses. Hot water tank standby losses represent heat that is lost through the tank walls to the surroundings. These losses are estimated to be 10% of the input energy into the tank [19–21]. Point-of-use losses represent heat that is given off to the surroundings when building occupants consume hot water at each point of use. Bertrand et al. conducted a review of point-of-use temperature losses that occur in common plumbing fixtures [9]. Most fixtures they reviewed had negligible point-of-use losses except for baths and showers. Baths are not common at Carleton University; however, showers are located in all campus residences and in the main athletics facilities. In the current study, a point-of-use energy loss of 12% is assumed due to shower use [9]. Sewer system losses represent heat that is lost through the sewer pipe walls into the surrounding earth as the effluent travels from the individual buildings to the main campus sewer outflow pipe. These losses are estimated to be 3% of the input energy [22, 23].

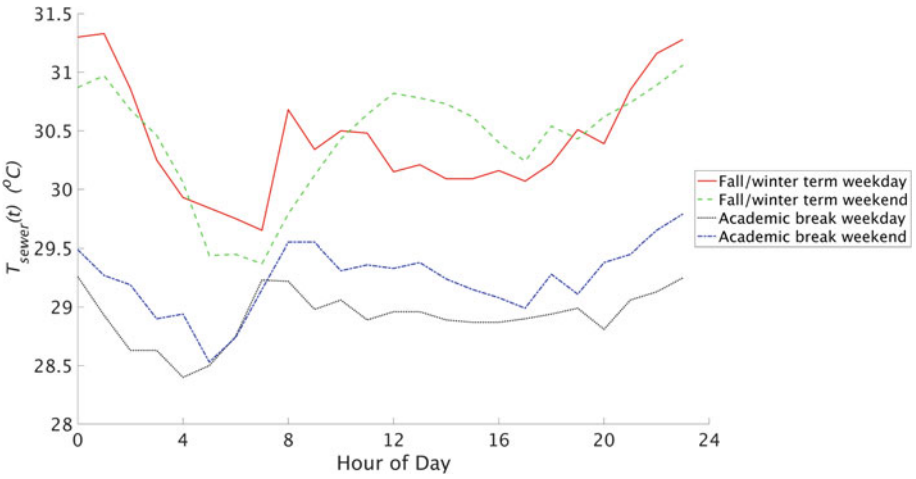


Fig. 3. Hourly variation of temperature, $T_{sewer}(t)$, in the main sewer outflow pipe over the course of an average day. Variations are shown for the same occupancy conditions depicted in Fig. 2. Hours 0–24 represent a 24-hour period from 12 am to 12 am.

The amount of energy that is recoverable from the main sewer outflow pipe for heating purposes is expressed as

$$\dot{Q}_{recoverable}(t) = \dot{V}_{sewer}(t)\rho C_p\left[T_{sewer}(t) - T_{discharge}\right] \tag{6}$$

Equation 6 is based on the assumption that the effluent leaving the campus (*i.e.*, the discharge) must always be at a temperature that is equal or greater than the average temperature of the mains water entering the campus. In the current study, the discharge temperature, $T_{discharge}$, is assumed to be constant at 4 °C. The amount of recoverable energy, given by Eq. 6, is shown in Fig. 4. For every data point in Fig. 4, an error bar is shown which corresponds to the maximum and minimum value of $\dot{Q}_{recoverable}(t)$ that is calculated when accounting for propagating parameter uncertainty in the analysis. The trend line linking the upper end of the error bars in each dataset in Fig. 4 represents the best-case recoverable energy (BCRE) estimation, whereas the trend line linking the lower end of the error bars represents the worst-case recoverable energy (WCRE) estimation. The parameters that were varied in the uncertainty analysis are listed in Table 1. Also shown in Table 1 are the ranges by which these parameters were varied to determine the BCRE and WCRE estimations.

2.2 Energy Models

Two energy models are constructed corresponding to a base case (BC) scenario and a sewer heat recovery (SHR) scenario. The BC scenario represents a base case natural gas boiler system that could be used at Carleton University for heating domestic hot water in the main athletics facility. The SHR scenario represents an alternative domestic hot water heating system consisting of a vapour-compression heat

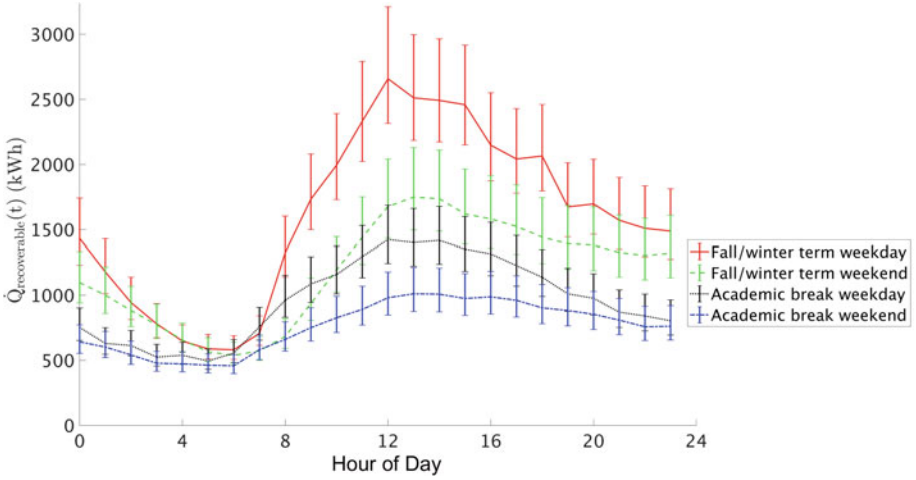


Fig. 4. Hourly variation of recoverable energy, $\dot{Q}_{recoverable}(t)$, in the main sewer outflow pipe over the course of an average day. Variations are shown for the same occupancy conditions that are depicted in Fig. 2. Hours 0–24 represent a 24-hour period from 12 am to 12 am.

pump-based sewer heat recovery system that is used to offset natural gas consumption by utilizing waste heat recovered from the main campus sewer outflow pipe. Both systems, shown schematically in Fig. 5, are sized to meet a peak domestic hot water heating load in the main athletics facility of 160 kW.

The BC system, shown in Fig. 5a, is assumed to be a high-efficiency natural gas condensing boiler. The energy output to meet the domestic hot water load in the BC system is given by

$$\dot{Q}_{out,BC}(t) = \dot{Q}_{fuel}(t)\eta_{BC} \tag{7}$$

where $\dot{Q}_{fuel}$ represents the energy content of the fuel entering the boiler, and η_{BC} is the average efficiency of the boiler, which is assumed to be equal to 95% [19, 30, 31].

The SHR system, shown in Fig. 5b, is assumed to be a vapour-compression heat pump-based sewer heat recovery system. When heat is required at the load, the effluent is pumped from the sewer wet well to the heat pump evaporator where energy is extracted from the fluid. A separator is located in between the wet well and the heat pump to discard solid waste from the effluent stream. The distance from the wet well to the main athletics facility is approximately 120 m. After leaving the heat pump, effluent mixes with the discarded solid waste and is returned to the wet well.

The energy required by the separator to operate internal pumps is a constant 0.9 kW [32]. The energy required to pump effluent from the wet well to the heat pump evaporator is expressed as

Table 1. Parameters and associated uncertainty ranges used to determine BCRE and WCRE estimations

Parameter	Assumed value	BCRE Value	WCRE Value	Parameter range rationale
Flow consumption	Measured value	−3% (of measured value)	+3% (of measured value)	Flow data varies ± 3% of its measured value, which is a typical range of measurement error in water flow metres [24, 25]
Steam consumption	Measured value	−5% (of measured value)	+5% (of measured value)	Steam energy consumption data varies ± 5% of its measured value, which is a typical range of measurement error in steam metres. [24, 26, 27]
Standby losses	10%	8%	12%	Standby losses vary by ±2% of assumed value to account for potential error in their estimation
Point-of-use losses	12%	10%	14%	Point-of-use losses vary by ±2% to account for potential error in their estimation
Sewer system losses	3%	1%	5%	Sewer system losses vary by ±2% to account for potential error in their estimation
Water mains temperature	4 °C	0.2 °C	10 °C	Water mains temperature varies from 0.2 °C to 10 °C to account for the seasonal variation of the water mains temperature. 0.2 °C and 10 °C represent the 2017 minimum and average annual measured wastewater treatment plant temperature in Ottawa, Ontario, respectively [28, 29]

$$\dot{W}_{in,ep}(t) = \frac{\dot{Q}_{Evap}(t)}{C_p[T_{sewer}(t) - T_{LLT}(t)]} g\Delta H \frac{1}{\eta_{ep}} \tag{8}$$

where $\dot{Q}_{Evap}(t)$ is the energy absorbed by the heat pump evaporator and $T_{LLT}(t)$ is the temperature at which the effluent leaves the evaporator. These parameters are taken from heat pump manufacture performance specifications [33, 34]. Also shown in Eq. 8 are the acceleration due to gravity, g, the effluent pump efficiency, η_{ep}, and the system head loss, expressed as

$$\Delta H = \left[\frac{fL}{d}\frac{v^2}{2g} + \sum K\frac{v^2}{2g}\right] \tag{9}$$

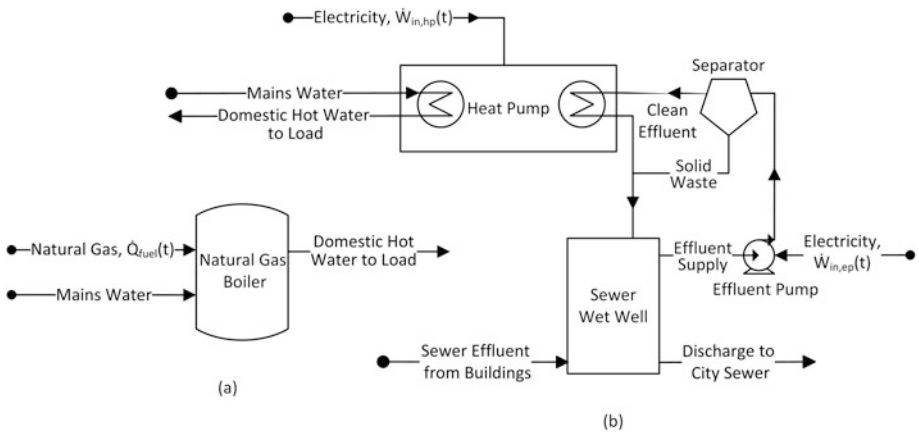


Fig. 5. a. BC system: high-efficiency natural gas condensing boiler. b. SHR system: vapour-compression heat pump-based sewer heat recovery system.

The parameters in Eq. 9 include the friction factor, f, the minor loss coefficient, K, the pipe diameter, d, the fluid velocity in the pipe, v, and the pipe length, L. The pipe diameter is determined using the following expression:

$$d = \sqrt{\frac{4\dot{V}_{\max,ep}}{\pi v_{econ}}} \tag{10}$$

where $\dot{V}_{\max,ep}$ is the maximum flow rate through the system and v_{econ} is the economic velocity of water. The values of several of the parameters shown in Eqs. 8–10 are listed in Table 2.

Table 2. Parameters used in determining the effluent pump energy requirement

Economic velocity, v_{econ}	Minor losses, K	Friction factor, f	Pumping distance, L	Effluent pump efficiency, η_{ep}	Effluent leaving temperature, $T_{LLT}(t)$
2 m/s[a]	30[a]	0.017[a]	240 m	0.84[b]	24–27 °C[c]

[a]Obtained from [35]
[b]Obtained from [36]
[c]Obtained from [33, 34]

The energy output to meet the domestic hot water load in the SHR heat pump system is expressed as

$$\dot{Q}_{out,SHR}(t) = \dot{W}_{in,hp}(t)COP(t) \tag{11}$$

where $\dot{W}_{in,hp}(t)$ represents the electricity input to drive the heat pump compressor, and $COP(t)$ represents the heat pump coefficient of performance (COP). Based on a first law, energy balance applied to the heat pump the COP is expressed as

$$COP(t) = 1 + \frac{\dot{V}_{sewer}(t)\rho C_p[T_{sewer}(t) - T_{mains}]}{\dot{W}_{in,hp}(t)} \tag{12}$$

In the current study, the heat pump COP varies roughly between 5.0 and 5.3, based on manufacturer specifications [33, 34].

The BC and SHR energy models are compared with respect to total annual cost and GHG emissions. Modelling parameters that are used to assess the total annual cost of each system include the capital cost, C_c, the annual maintenance cost, M_c, the variable fuel cost, F_c, the fuel escalation rate, k, and the system lifetime, n, as shown in Table 3. A discount rate, i, of 8% is applied to both systems [37]. The total annual cost (AC) is expressed as

$$AC = NPC\left[\frac{i(1+i)^n}{(1+i)^n - 1}\right] \tag{13}$$

where the net present cost (NPC) is given by

$$NPC = C_c + \sum_j \frac{M_c + F_c}{(1+i)^j} \tag{14}$$

The net present cost is calculated over j annual periods in which costs are incurred.

Table 3. Modelling parameters considered in feasibility assessment of heat pump-based sewer heat recovery system. All costs shown are in Canadian dollars.

System	Capital cost, C_c (k$)	Annual maintenance cost, M_c ($/year)	Variable fuel cost, F_c	Fuel escalation rate, k (%/year)	Lifetime, n (years)	Fuel emission factor, e_{CO_2e}
Heat pump-based sewer heat recovery system	76[a]	210[b]	120[c] ($/MWh)	6[d]	20[f]	0.043 [g] (kg CO_2e / MWh)
Natural gas boiler system	31[b]	290[b]	0.29[c] ($/m^3)	2[e]	13[b]	186 [h] (kg CO_2e/MWh)

[a]Obtained from [36, 38–41]. It includes costs to purchase and install pipe between the sewer outflow pipe and main athletics facility
[b]Obtained from [40]
[c]Obtained from [42]
[d]Obtained from [43]
[e]Obtained from [44]
[f]Obtained from [45, 46]
[g]Obtained from [47]
[h]Obtained from [48, 49]

Also shown in Table 3 are the corresponding fuel GHG emission factors, e_{CO_2e}, for natural gas, and electricity generation on the Ontario grid. These factors are used to

calculate the associated GHG emissions in equivalent metric tonnes of CO_2, $m_{CO_{2e}}$, resulting from the implementation of each system, as shown in Eq. 15.

$$m_{CO_{2e}} = \frac{AEO(e_{CO_{2e}})}{1000} \tag{15}$$

The variable *AEO* in Eq. 15 represents the annual energy output, in units of MWh, of each system shown in Table 3.

2.3 Validation

The total monthly building-based water consumption data used in this study is validated experimentally using measured data from the three water mains that supply the university campus. The district heating grid condensate flow profile and the building flow profiles are combined and compared to the monthly measured data from the three water mains from the past 2 years. The former is calculated to be 1.2% lower on average than the latter. Potential reasons for this discrepancy can be attributed to leaks in the distribution system between the city water metres and the building water metres, and/or landscape irrigation that occurs in the summer. Overall, this comparison indicates a close match between the datasets and validates our estimate for the volume of water flowing into the sewer.

A validation of our energy use estimate for domestic hot water heating in each building on campus is also performed. The validation is accomplished by comparing estimated domestic hot water consumption values for various building types on campus with values published in the literature, as shown in Table 4. Since standard published data for building domestic hot water use is typically available on a daily and per occupant basis, a method outlined by Abdelalim et al. [42] is used to convert our estimated data into a similar format. Using this method, building occupancy schedules are determined for the various building types located on the Carleton University campus. From these schedules, the daily domestic hot water consumption per occupant is calculated.

Table 4 shows that the estimated domestic hot water consumption values for residence, office, and classroom buildings are below the published values. One potential reason for this discrepancy is due to low flow fixtures being installed in most buildings fitting this description on the Carleton campus. The percent difference is especially high for residence buildings, since these buildings have more washroom facilities and as a result consume more domestic hot water. Domestic hot water consumption estimates for cafeteria buildings, on the other hand, are shown to be considerably greater than the published values. One reason for this may be that published studies typically focus on restaurants or cafeterias that have one to two meal periods over an average workday [16]. However, many cafeterias at Carleton University have three meal periods in a day and are commonly open until late at night [52]. Having extended hours of operation in these cafeterias would likely increase domestic hot water consumption above standard published values. As no published domestic hot water consumption data is available for laboratory and sports facilities, a comparison cannot be made with our estimated values. These buildings are very diverse and contain specific equipment that is not

Table 4. Carleton building hot water use per person compared with published values for validation

Building types	Estimated domestic hot water consumption (l/day/person)	Published domestic hot water consumption (l/day/person)	Percent difference (relative to published domestic hot water consumption)
Residences	34.4	48[a]	−28%
Office and classroom	3.6	3.8[a]	−5%
Mixed use with a large cafeteria	28.8	14[b]	106%
Sport facilities	26.5	N/A	N/A
Laboratory buildings	69.2	N/A	N/A

[a]Obtained from [50, 51]
[b]Average value obtained from [16]

commonly found outside of a university campus. Carleton's library building is also excluded from this analysis due to the challenges associated with estimating the building's occupancy.

3 Results and Discussion

The amount of heat that can potentially be recovered on an annual basis from the main campus sewer outflow pipe is estimated to be 10 850 MWh. The hourly maximum and minimum amounts of recoverable heat over the course of the year are estimated to be 2660 kWh and 458 kWh, respectively, as depicted by the daily occupancy profiles in Fig. 4. In comparison, the total annual and peak hourly domestic hot water loads in the main athletics facility are relatively small at 730 MWh and 160 kWh, respectively. Thus, in order to meet this annual domestic hot water load, only a fraction of the total recoverable sewer heat shown in Fig. 4 is required. Furthermore, as the heat pump in the SHR scenario is sized to meet the peak load, no auxiliary heating system (e.g., natural gas boiler) is required for peaking purposes.

The total annual cost and annual GHG emissions corresponding to the BC scenario are calculated as $30 900 and 143 metric tonnes of CO_2e, respectively. Figure 6 shows the percentage by which these parameters are reduced relative to the BC scenario as a result of implementing the SHR scenario. The error bars in Fig. 6 show the variation in percent reduction for both the BCRE and WCRE estimations, where the upper end of the error bars represents the BCRE estimation, and the lower end of the error bars represents the WCRE estimation. Two SHR scenario configurations are shown in Fig. 6: one that includes district heating grid condensate, and one that does not. The total annual cost decreases by approximately 8% in the SHR scenario (with condensate) relative to the BC scenario, which is equivalent to $2550 annually. The annual GHG

emissions decrease by approximately 95% in the SHR scenario (with condensate) relative to the BC scenario, which is equivalent to 136 metric tonnes of CO_2e annually. GHG emissions in the SHR scenario are mainly due to electricity usage in both the effluent pump and heat pump compressor. The high GHG emission reductions associated with this scenario demonstrate the potential that sewer heat recovery systems have with respect to lowering emissions in our energy system.

Also shown in Fig. 6 are the percent reductions of the same parameters mentioned above, but for the SHR scenario (without condensate). This hypothetical system is analysed to assess the impacts of implementing a heat pump sewer heat recovery system at the university campus if it did not contain a steam-based district heating grid. The average temperature in Carleton's sewer system is approximately 30 °C, which is above what has been typically documented in previous studies [7]. The reason for this high temperature is due to the condensate that enters the sewer system from the district heating grid. The condensate does not cool before entering the sewer and therefore carries a significant amount of thermal energy, which raises the overall temperature of the sewer flow. In this study, the condensate from the steam district heating grid raises the temperature in the sewer by an average of 5.7 °C relative to a sewer system without condensate. Without the condensate from the district heating grid, the amount of heat that can potentially be recovered on an annual basis from the main campus sewer

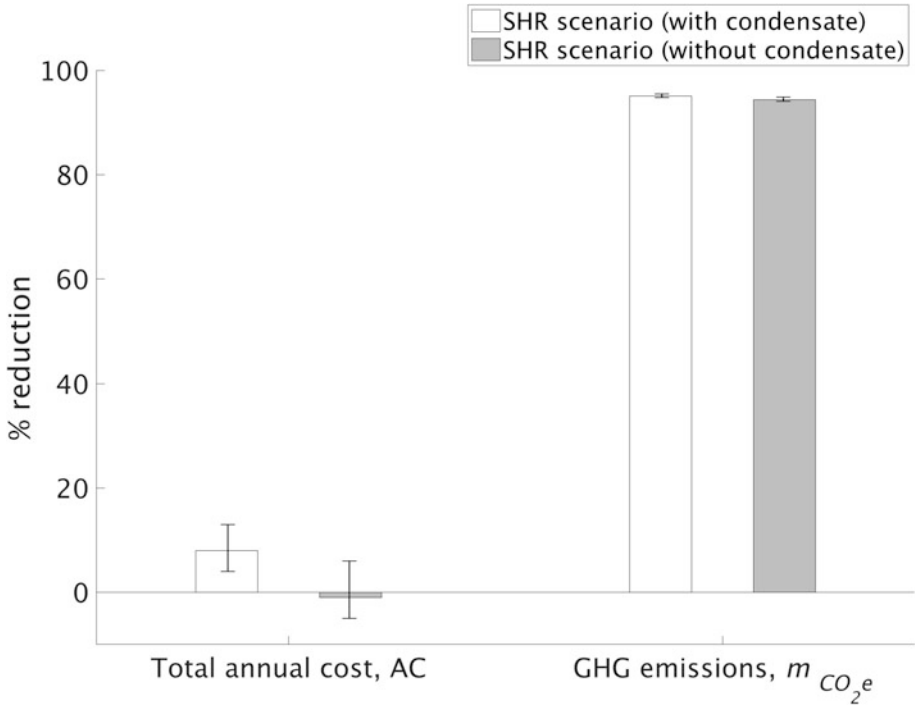


Fig. 6. Percent reduction in annual cost and GHG emissions relative to the BC scenario as a result of implementing the SHR scenario (with and without condensate)

outflow pipe is estimated to be 8470 MWh. This corresponds to an annual decrease in the recoverable heat of approximately 22% relative to a sewer system with condensate. As in the SHR scenario (with condensate), the heat pump is sized to meet the peak load; therefore, no auxiliary heating system is required for peaking purposes.

Figure 6 shows the percentage by which the total annual cost and annual GHG emissions are reduced in the hypothetical system (i.e., without condensate) as a result of implementing the SHR scenario. The total annual cost increases by approximately 1% in the SHR scenario (without condensate) relative to the BC scenario, which is equivalent to $370 annually. This cost increase is mainly due to inefficiencies associated with the sewer heat recovery system's performance. As the effluent temperature decreases, the heat pump COP also decreases, resulting in increased electricity use in the compressor and increased operating costs. Moreover, higher flow rates are required to extract an equivalent amount of thermal energy from a lower temperature source. Therefore, electricity use in the effluent pump is also increased, resulting in increased operating costs. The annual GHG emissions decrease by approximately 94% in the SHR scenario (without condensate) relative to the BC scenario, which is equivalent to 135 metric tons of CO_2e annually. Despite the setback of having a small cost increase, the SHR scenario (without condensate) shows considerable promise with regard to lowering GHG emissions.

A number of parameters are identified in the current study that impact the economic feasibility of the SHR scenario. Figure 7 shows the sensitivity of the total annual cost savings in the SHR scenario (with condensate) with respect to a variety of system parameters such as the effluent temperature, lifetime, capital cost, fuel escalation rate, discount rate, and distance between load and main sewer outflow pipe. The impacts shown in Fig. 7 are calculated by varying these system parameters by 30% of their estimated values.

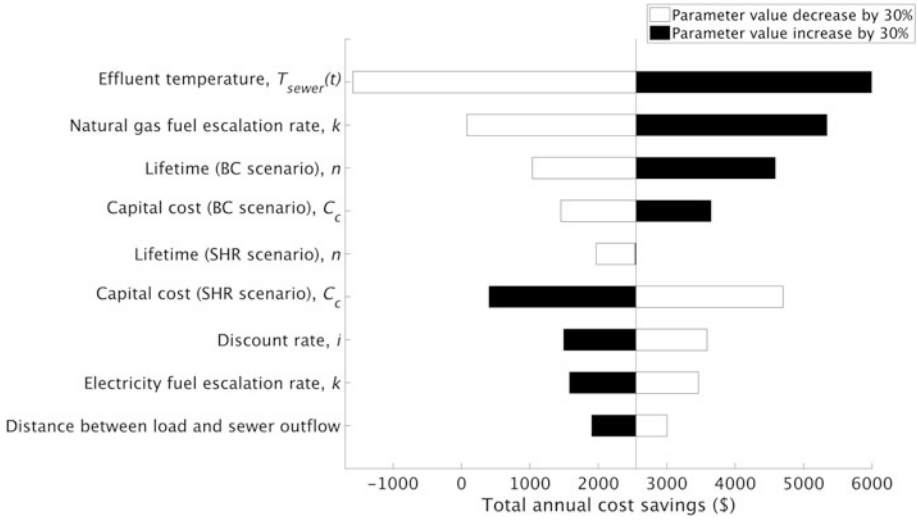


Fig. 7. Sensitivity of total annual cost savings in SHR scenario (with condensate) with respect to a number of system parameters

Figure 7 shows that increasing the effluent temperature, natural gas fuel escalation rate, BC scenario lifetime, and BC scenario capital cost by 30% causes the total annual cost savings to increase by $3450, $2790, $2040, and $1100, respectively. However, increasing SHR scenario lifetime, SHR scenario capital cost, discount rate, electricity fuel escalation rate, and the distance between the load and the main sewer outflow pipe by 30% causes the total annual cost savings to decrease by $10, $2150, $1060, $970, and $650, respectively. On the other hand, decreasing each of these system parameters by 30% causes the total annual cost savings to change by an amount that is approximately equal and opposite to those mentioned above, as shown in Fig. 7. The only exception to this applies to the SHR scenario lifetime. Decreasing this parameter by 30% causes the total annual cost savings to decrease by $580. The most sensitive parameter in Fig. 7 is the effluent temperature, followed by the natural gas fuel escalation rate, and the SHR scenario capital cost. Efforts should be made to lower the uncertainty in estimating these parameters as they have the greatest impact on the total annual system cost. Another parameter that shows less of an impact on total annual cost savings in Fig. 7 is the distance between the load and the main sewer outflow pipe. This result can be misleading, as it is highly dependent on the estimated distance between the load and the main sewer outflow pipe. In the current study, this distance is equivalent to 120 m, thus a change of 30% only represents an addition or removal of 40 m of piping between the two points. In the case that the estimated distance would be 1000 m, the impact on the total annual cost of the system would be considerably higher. The sensitivity of this parameter is therefore highly geographically dependent.

A number of other parameters not mentioned in this study are capable of having a significant impact on the economic feasibility of the SHR scenario. One example is the adverse impacts related to heat exchanger fouling in sewage heat recovery systems. Heat exchanger fouling is common in these systems and is highly dependent on the chemical and biological composition, as well as the temperature of the effluent stream [7, 53]. Given that the temperature of the effluent stream at Carleton University is higher than average due to the addition of condensate from the district heating grid, there is a high likelihood that heat exchanger fouling would be an issue. However, further detailed component analysis would be required to make a more complete estimate of fouling effects. Another measure that would affect the feasibility of the SHR scenario is the implementation of a carbon tax. A carbon tax is not currently considered in the analysis; however, Canada has committed to implementing a carbon tax of 50 $/ton by the year 2022 [54]. This level of carbon tax would provide an economic benefit to the large GHG emission reductions occurring in the SHR scenario, and potentially improve the economic feasibility of the system.

The current study shows that sewer heat recovery systems are an economically viable option for university campuses that derive a significant portion of their heat from a steam-based district heating grid. These systems are also capable of significantly reducing GHG emissions relative to a conventional natural gas-fired boiler system. Sewer heat recovery systems will likely play an important role in the future heating market as fossil fuel costs increase and GHG emission reductions become a higher priority.

4 Conclusions

The annual sewer waste heat resource at Carleton University was estimated using a novel method. This method allows for the sewer temperature and flow rate to be estimated based on readily available building metered data such as water and energy consumption. A model of the university energy system was constructed, and two scenarios were developed and compared with respect to total annual cost and annual GHG emissions. In the first scenario, the existing natural gas boiler system is used to meet the domestic hot water load in the main athletics facility. In the second scenario, this load is met using a heat pump-based sewer heat recovery system. Results show that switching to the latter system causes the total annual cost and annual CO_2 emissions to decrease by 8% and 95%, respectively.

Although these results sound promising, it should be noted that sewer flows at this campus are warmer than average due to the addition of steam condensate from the district heating grid, resulting in ideal conditions for waste heat recovery. When the steam condensate input is removed from the sewer effluent flow, the total annual cost increases by 10% relative to the natural gas boiler system, and CO_2 emission reductions remain approximately the same. This minor cost setback is mainly believed to be due to the size of the sewer heat recovery system that is used. The sewer heat recovery system assessed in the current study is small in scale and therefore makes use of only a fraction (i.e., roughly 7%) of the recoverable heat that is available in the sewer. A larger scale system would likely result in greater annual cost reductions due to economies of scale, with or without condensate addition from a district heating grid.

Implementing a sewer heat recovery system at the Carleton University campus has been found to be a viable option from both an economic and environmental standpoint. Based on these findings, one can presume that other campuses and municipalities with comparable building stock and system conditions located across North America could benefit from this technology. In the case that sewer temperature and flow data are not readily available in some of these areas, the methods proposed in this study may be used as a general guideline for estimating the recoverable sewer heat potential.

Acknowledgements. The authors would like to thank the Carleton Discovery Center for partly funding this research through their ICUREUS programme. The authors would also like to thank Carleton University's Facilities Management and Planning group for their close cooperation and support in acquiring building consumption data for this project.

References

1. Government of Canada: Canada's 2017 Nationally Determined Contribution submission to the United Nations Framework (2017)
2. Environment and Climate Change Canada: Canada's 2016 Greenhouse Gas Emissions Reference Case. Canada (2017)
3. Jordaan, S.M., Romo-Rabago, E., McLeary, R., Reidy, L., Nazari, J., Herremans, I.M.: The role of energy technology innovation in reducing greenhouse gas emissions: a case study of

Canada. Renew. Sustain. Energy Rev. **78**, 1397–1409 (2017). https://doi.org/10.1016/j.rser.2017.05.162
4. Natural Resources Canada: Energy Fact Book. Natural Resources Canada, Canada (2016)
5. Environment and Climate Change Canada: Canadian Environmental Sustainability Indicators Greenhouse Gas Emissions. Canada (2017). https://doi.org/10.1016/b978-0-12-409548-9.05178-2
6. Natural Resources Canada: Improving Energy Performance in Canada. Natural Resources Canada, Canada (2015)
7. Hepbasli, A., Biyik, E., Ekren, O., Gunerhan, H., Araz, M.: A key review of wastewater source heat pump (WWSHP) systems. Energy Convers. Manag. **88**, 700–722 (2014). https://doi.org/10.1016/j.enconman.2014.08.065
8. Cipolla, S.S., Maglionico, M.: Heat recovery from urban wastewater: analysis of the variability of flow rate and temperature in the sewer of Bologna, Italy. Energy Proc. **45**, 288–297 (2014). https://doi.org/10.1016/j.egypro.2014.01.031
9. Bertrand, A., Aggoune, R., Maréchal, F.: In-building waste water heat recovery: an urban-scale method for the characterisation of water streams and the assessment of energy savings and costs. Appl. Energy **203**, 983 (2017). https://doi.org/10.1016/j.apenergy.2017.07.106
10. Wallin, J., Claesson, J.: Investigating the efficiency of a vertical inline drain water heat recovery heat exchanger in a system boosted with a heat pump. Energy Build **80**, 7–16 (2014). https://doi.org/10.1016/j.enbuild.2014.05.003
11. Dong, J., Zhang, Z., Yao, Y., Jiang, Y., Lei, B.: Experimental performance evaluation of a novel heat pump water heater assisted with shower drain water. Appl. Energy **154**, 842–850 (2015). https://doi.org/10.1016/j.apenergy.2015.05.044
12. City of Vancouver: Neighbourhood Energy Utility. City of Vancouver, Canada (2012)
13. FVB Energy.: Southeast False Creek Neighbourhood Energy Utility. http://www.fvbenergy.com/projects/southeast-false-creek-neighbourhood-energy-utility/ (2018). Accessed 2 May 2018
14. Im, P., Liu, X., Henderson, H.: Operational performance characterization of a heat pump system utilizing recycled water as heat sink and heat source in a cool and dry climate. Energies **11**, 211 (2018). https://doi.org/10.3390/en11010211
15. Office of Institutional Research and Planning: Facts and Figures Carleton University. https://carleton.ca/about/facts/ (2017). Accessed 1 May 2018
16. Fuentes, E., Arce, L., Salom, J.: A review of domestic hot water consumption profiles for application in systems and buildings energy performance analysis. Renew. Sustain. Energy Rev. **81**, 1530–1547 (2018). https://doi.org/10.1016/j.rser.2017.05.229
17. Symonds, G.: Carleton Facilites Management and Planning, Boiler Make-up Water. Personal Communication (2018)
18. Macdonald, S.: Carleton Facilities Management and Planning, Waste Heat Recovery Project. Personal Communication (2018)
19. Intertek Testing Services: Consumers' Directory of Certified Efficiency Ratings for Heating and Water Heating Equipment. Intertek Testing Services, New York (2008)
20. Biermayer P, Lutz J.: Gas water heater energy losses. EScholarship, 1–23 (2009)
21. Aguilar, C., White, D.J., Ryan, D.L.: Domestic water heating and water heater energy consumption in Canada (2005)
22. Dürrenmatt, D.J., Wanner, O.: A mathematical model to predict the effect of heat recovery on the wastewater temperature in sewers. Water Res. **48**, 548–558 (2014). https://doi.org/10.1016/j.watres.2013.10.017
23. International Renewable Energy Agency: Renewable Energy in District Heating and Cooling, A Sector Roadmap for REmap. UAE, Abu Dhabi (2017)

24. Parker, S.A., Boyd, B.K., Hunt, W.D., Stoughton, K.L.M., Fowler, K.M., Koehler, T.M. et al: Metering Best Practices: A Guide to Achieving Utility Resource Efficiency, Release 3.0. (2015)
25. Barfuss, S., Johnson, M., Neilsen, M.: Accuracy of In-Service Water Meters at Low and High Flow Rates. Logan, USA (2011)
26. Gruskos, R.: Utility Flow Metering for Steam and Heated/Chilled Water: A Tutorial. International District Energy Association (2015)
27. Szyszko, G.: Energy and Flow Measurement for Hydronic Systems. Association of Energy Engineers Southern California (2013)
28. City Of Ottawa: Lemieux Water Purification Plant—2017 Drinking Water Quality. Canada (2017)
29. City Of Ottawa: Britannia Water Purification Plant—2017 Drinking Water Quality. Canada (2017)
30. Office of Energy Efficiency & Renewable Energy: Purchasing Energy-Efficient Commercial Boilers. Washington DC, USA (2016)
31. Lunt, J.: Condensing storage water heaters. J Light Constr. 1–7 (2009)
32. SHARC Energy Systems: Technical Information: SHARC 660SS Wastewater Solids Filter. Canada (2015)
33. Maritime Geothermal: Engineering Specification: High Temperature Water to Water Heat Pump. New Brunswick (2018)
34. Mitsubishi Heavy Industries: Water to Water Centrifugal Heat Pump; Heat Recovery Type. Japan (2018)
35. Janna, W.S.: Design of Fluid Thermal Systems, Third. Cengage Learning, Stamford, USA (2015)
36. Cat Pumps: Pumps and Accessories Price List. Minneapolis, USA (2012)
37. Jenkins, G., Kuo, C-Y.: The economic opportunity cost of capital for Canada—An empirical update. Discount Rates Eval. Public Priv. Partnerships, Kingston, Canada (2008)
38. State of Michigan: Miscellaneous Industrial Costs, vol. 12. Lansing, USA (2003)
39. Thomson, J.A.: National Plumbing & HVAC Estimator. Associated General Contractors of America, Carlsbad, USA (2016)
40. US Energy Information Administration: Updated Buildings Sector Appliance and Equipment Costs and Efficiency (2016)
41. Duquette, J.: Carleton University, Cost Data. Personal Communication (2018)
42. Abdelalim, A., O'Brien, W., Shi, Z.: Visualization of energy and water consumption and GHG emissions: a case study of a Canadian University Campus. Energy Build **109**, 334–352 (2015). https://doi.org/10.1016/j.enbuild.2015.09.058
43. Ontario Ministry of Energy: Ontario's Long Term Energy Plan: Delivering Fairness and Choice. Ontario, Toronto (2017)
44. World Bank: Commodity Markets Outlook—World Bank. Washington DC, USA (2017)
45. SHARC Energy Systems: Case Study 1 Sharc System—Leisure Centre. Nottingham (2014)
46. United States Department of Energy: Geothermal Heat Pumps 2018, pp. 10–3. https://www.energy.gov/energysaver/heat-and-cool/heat-pump-systems/geothermal-heat-pumps (2018). Accessed 11 April 2018
47. The Atmospheric Fund: A clearer view on Ontario's emissions: practice guidelines for electricity emissions factors (2017)
48. Office of Air Quality Planning and Standards AP-42: Compilation of Air Pollutant Emission Factors. North Carolina (1995)
49. US Environmental Protection Agency: Overview of Greenhouse Gases 2018, p. 5. https://www.epa.gov/ghgemissions/overview-greenhouse-gases#methane (2018). Accessed 12 April 2018

50. ASHRAE: ASHRAE Handbook—Service water heating. ASHRAE (2015)
51. Dewerth, D., Becker, B.: Comparison of collected and compiled existing data on service hot water use patterns in residential and commercial establishments (1993)
52. Carleton University: Carleton Dining Services 2018. https://dining.carleton.ca/ (2018). Accessed 2 May 2018
53. Culha, O., Gunerhan, H., Biyik, E., Ekren, O., Hepbasli, A.: Heat exchanger applications in wastewater source heat pumps for buildings: a key review. Energy Build **104**, 215–232 (2015). https://doi.org/10.1016/j.enbuild.2015.07.013
54. Environment and Climate Change Canada: Technical Paper on the Federal Carbon Pricing Backstop 2017, p. 26 (2017)

Sustainability and Energy Services: A Framework for Discussion

John H. Perkins(✉)

The Evergreen State College, 236 Cambridge Avenue, Kensington, CA 94708, USA
perkinsj@evergreen.edu

Abstract. Current energy services provide incalculable benefits, but they are not sustainable. Several factors cause unsustainability, but climate change is the most important and immediate risk. Moving to new energy systems requires focus on investment: the steady stream of monetary and social capital to build, maintain, operate, regulate, and decommission energy infrastructure. Investments develop primary energy sources and make them available to societies as energy services. Almost without exception, an investment must focus on a fuel–technology pair. An investment to develop natural gas, for example, cannot yield PV electricity, nor can an investment in wind turbines make coal into a useful source. Only nine primary energy sources are usable for energy services, and widespread assumptions hold that investments in renewable energy are more sustainable than investments in the currently dominant nonrenewable sources. Moving to higher dependence on renewable energy, however, requires the development of a finer analysis of the sustainability of the nine different energy sources. This paper proposes an energy–sustainability matrix, composed of five criteria currently in use, supplemented by four new criteria drawn from the concept of sustainability. Qualitative, ordinal standards then enable ranking of each of the nine primary energy sources, plus energy efficiency, with the nine criteria for investments. A complete analysis indicates the comparative strengths and weaknesses of each energy source. This matrix usefully frames a sociopolitical discussion for public policy and investment decisions.

1 Introduction

It is easy, true, and trivial to say that energy is an indispensable element of modern, industrial societies. Everybody knows this fact, and governments around the world gear their policies toward protecting and stabilizing supplies of primary energy sources and delivery of electric power. Political leaders who through action or inaction fail to protect energy supplies will likely find themselves no longer in power.

Despite the obvious importance of energy and energy services, it has proven quite difficult or impossible to address the full range of issues concerning energy. These difficulties pervade debates and discussions in scientific and technical circles, the

Member of the Faculty *Emeritus*.

A. Vasel and D. S.-K. Ting (Eds.): EAS 2018, SPE, pp. 168–181, 2019.
https://doi.org/10.1007/978-3-030-00105-6_10

political arena, business and economic circles, and everyday conversations. Such difficulties are not surprising given that energy, energy supplies, and energy services [1] involve complicated technical issues, deep connections to both economic and political power, and an array of ethical and moral questions that can stymie even the most clear-headed of thinkers.

Dissent, disagreement, and conflict accordingly surround energy conversations, but climate change makes the incoherence of energy conversations dangerous and unacceptable. In a word, climate change makes the world's current energy economy unsustainable: 81% of energy services currently rely on fossil fuels (coal, oil, and gas), the major cause of climate change [2]. The immense benefits of these primary energy sources are not in doubt, but climate change poses potentially catastrophic risks to the modern civilizations made possible by them. To avoid these risks, it is imperative to resolve energy arguments and agree on workable plans for transition to more sustainable energy services.

This paper proposes an analytical framework based on continual investments needed to procure, refine, transport, maintain, and use primary energy sources. Historical analysis suggests that investors have used five criteria to decide investments, but sustainability concerns now demand four additional criteria to join with the original five. Criteria identify goals, and standards assess the strengths and weaknesses with which energy sources satisfy each criterion. Ordinal, qualitative standards enable ranking of competing investment proposals. This assessment framework differs from but overlaps other theoretical approaches focused on environmental justice and poverty [3], the political economy of decision-making processes about energy transitions [4], and goals for sustainable development [5], but comparing the four takes us too far afield for this paper.

2 Energy, Investment, and Sustainability

Homo sapiens may be at the top of the list of a few species in traits like language and toolmaking, but it is alone in its abilities to manipulate combustion of firewood—the primeval source of external energy. Where all other species avoid or flee fire, and never control it, people congregated around and used energy to shape their societies and lives. Human abilities to control fire and their dependence on it probably originated about one million years ago, and cooked food, one of fire's first uses, may have played an important role in the evolution of *Homo sapiens* [6]. Now, nine sources of heat or electromotive force comprise the primary energy resources: coal, oil, gas, uranium, falling water, solar radiation, wind, biomass, and geothermal heat.

2.1 Energy

Despite the ancient centrality of external energy to human life, the concept of energy with a precise meaning has a history of only about two centuries. The ideas behind the English word *energy* emerged among European artisans, engineers, and scientists in the 1800s and acquired its first, precise, mathematical meaning from the works of Carnot, Clapeyron, Faraday, Mayer, Joule, Thomson, Clausius, Maxwell, and others. Energy

united wide-ranging, rapidly evolving work and study with steam engines, heat, and electricity. Today, the idea of energy in its various forms plays a central role in all engineering and natural science, and its behavior is codified in the Laws of Thermodynamics [7].

Energy's entrance to engineering and science is comparatively recent, but the entrance of primary energy sources (fuels) into commerce and the exercise of political power dates to ancient times. None of the nine primary energy sources are usable for free, perhaps other than using sunlight to see or turning to face a warming sun on a cold day. Practical uses of an energy source require humans to expend money and effort to gather, process, and harvest the energy. In turn, expenditures of money and effort have always conferred value on energy efficiency.

2.2 Investment

Consider the first uses of firewood (now called biomass): early humans found it easily available in abundant quantities, but to make use of it, the wood had to be gathered and hauled to the human settlement. There it might have to be broken into smaller pieces and dried. Concentrating the heat of combustion required hearths and cooking utensils. Once burned, the ashes required physical removal to prepare for the next batch. The fuel may have been free and available for the taking, but using it was not. People had to invest work and energy, derived from food to power their muscles, to make firewood's energy available for tasks like cooking and keeping a shelter lighted and warm. Using energy was never a "free lunch," and today investment in energy is the *sine qua non* for producing energy services. It is very simple: no investment, no energy service.

Intrinsic with investment was the element of choice, because even early humans could elect to acquire firewood in different ways. For example, people undoubtedly discovered early in their relationship with fire that wood from some trees gave more heat than other kinds of trees, per unit of effort in gathering and preparing the wood for use. They also undoubtedly learned early on that dry wood burned faster and hotter than green, wet wood. The pleasures of firewood's energy services—heat and light—required effort to obtain, but why spend any more effort than necessary to obtain these benefits?

If we accept the existing estimate that human dependence on fire began about one million years ago, we can divide that time into 1,000 millennia. For about 999.5 of those millennia, firewood was either the only or certainly the most important source of energy, other than food. In a few areas, coal outcroppings, oil seeps, wind, falling water, or geothermal hot springs may have provided small, additional quantities of energy services. Only in the last 0.5 millennium (500 years) have more primary energy sources become available in significant quantities.

Today, human societies draw upon nine primary energy sources, but investment has become much more complicated. Technical and physical feasibility intertwines with considerations of economic, sociopolitical, and geographic constraints to make a nexus of issues with compelling ethical and moral questions. In addition, previous investments leave a legacy of infrastructure, customs, and human relationships, which shape future decision-making. Which primary energy source will best provide energy services

in quantities and at costs people find acceptable, given existing political economic conditions?

2.3 Sustainability

Investment choices to develop primary energy sources emphasized the fossil fuels (coal, oil, and gas) from the sixteenth century forward. If the uses of these materials could continue indefinitely, providing many benefits and no significant damages, then no problem of sustainability would exist. Unfortunately, that happy situation does not exist.

Health and environmental effects plague certain areas due to combustion wastes, waste disposal, and mining disruptions. At various times, geopolitical tensions have erupted into conflict and war, especially over oil. If these were not damaging enough, the remaining supplies of fossil fuels are finite. Deposits remaining in the earth can support continued uses for many years, but ultimately further mining efforts will be useless [8].

Climate change, however, tops all concerns about fossil fuels. Burning these materials for heat and light releases the greenhouse gas carbon dioxide (CO_2) and has increased its concentration in the atmosphere. Methane (CH_4), the primary constituent of natural gas, is another greenhouse gas. Methane is released into the atmosphere via several sources, including energy production. As greenhouse gases trapping heat, CO_2 and CH_4 account for most of the temperature increases observed since 1750 [9].

Climate change and the other problems make continued reliance on fossil fuels unsustainable. Sustainability is a wonderful word, partly because it sounds so comforting and partly because it is flexible in its common usages. Unfortunately, it is so flexible that it all too frequently has no clear meaning. This article uses the most common meaning of the word and supplements it with specific attributes.

In 1986, the World Commission on Environment and Development, organized by the United Nations and chaired by Gro Harlem Brundtland, Norway's former prime minister, released *Our Common Future*, which defined sustainability as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs." [10]

Precision of the word comes from these questions [11], with my responses in (parentheses). Note that energy services—which are to be sustained—are neutral about the energy source providing the service:

- What is to be sustained? (Energy services),
- Over what area? (Every country),
- Over what time? (Forever, or a very long time),
- At what cost? (Reasonable), and
- For whose benefit? (Everyone, now and in the future).

3 The Energy–Sustainability Matrix

Sustainability opens the pathway to a new type of systematic consideration of the strengths and weaknesses of energy sources in providing energy services. A matrix is formed (Table 1) by making the ten energy sources (nine primary energy sources plus energy efficiency) a column. On the horizontal axis, nine criteria—described below—specify desired goals that an individual, community, company, society, nation-state, or the world wishes to obtain or achieve. The energy–sustainability matrix (hereafter, the "matrix") I propose thus has 90 cells ($10 \times 9 = 90$) awaiting analysis to indicate a standard: how well a source satisfies a criterion.

An important note about the matrix: In Table 1, all criteria appear to be of equal rank. In assessments, however, some will rank certain criteria higher, perhaps far higher, than others. For example, investors seeking profits may value Profits and Risks much higher than Protection of Common Resources or Ethics. Similarly, an environmental analyst may value Protection of Common Resources much higher than Size of Resource or Aesthetics.

Table 1. The energy–sustainability matrix

Criteria	Profits and risks	Security of access and profits	Size of resource	Time of use	Geographic distribution	Protection of common resources	Distribution of benefits, costs, and risks	Ethics	Aesthetics
Source	Traditional criteria					Sustainability criteria			
Coal									
Petroleum									
Gas									
Uranium									
Efficiency									
Solar									
Wind									
Hydropower									
Biomass									
Geothermal									

3.1 Traditional Criteria

Historical analysis of the world's energy economy for the past 500 years suggests that five criteria guided individuals, communities, companies, and governments as they invested and then promoted and protected their energy enterprises and infrastructures [12].

Profits and Risks. As human societies developed, especially in the last five to ten millennia, division of labor and the use of money for commercial exchanges became increasingly common, and fuels became a commercial product. Some people gathered wood, hauled it to cities, and sold it to people who did not gather wood. In sixteenth century England, near-surface coal deposits yielded coal to those who mined it and sold it to people who did not mine coal. Coal increasingly replaced firewood starting in

sixteenth century England, because supplies of firewood became increasingly scarce, especially around growing cities like London [13].

By the end of the nineteenth century, investors increasingly mined coal, pumped oil, captured natural gas from wells, and marshaled falling water to make electricity. A few countries, mostly in Europe, the United States, and Japan, developed substantial industries to make and sell energy. These companies procured, processed, and transported energy to customers; other companies sold new devices, such as light bulbs and automobiles, so that energy customers could enjoy energy services, such as light and mobility.

One overarching rule governed these transactions: the merchants selling energy had to receive at least as much money as it costs them to deliver their products. In short, they had to break even to stay in business, and of course the aim of energy purveyors was to profit above expenses. Moreover, the risks they incurred had to be manageable or tolerable if they were to remain in the game. In the modern nation-state, energy became a major business—subject to government rules, taxes, and subsidies—and the profits and risks of the business had to look acceptable to attract investors.

Security of Access and Profits. This criterion is a subset of *profits and risks*, but it became of distinct importance during the last half of the nineteenth century and especially in the twentieth and twenty-first centuries. Energy businesses became large, integrated operations that often sought to procure energy sources in foreign countries. Procurement abroad brought heightened risks of political instability that a company's home country could not control. For example, the launch of the Anglo-Persian Oil Company (now BP) by the United Kingdom in 1909 soon required the British government to acquire 51% of the company's stock, an expense Parliament undertook partly to assure a secure source of oil for the Royal Navy [14]. Security of access and profits took on imperial dimensions as the world's energy economy expanded rapidly in the twentieth century, and investors insisted that their access to physical supplies be safe and that any profits earned would belong to the investors.

Size of the Resource. Like most businesses, energy companies faced fixed and variable costs. For example, developing a coal mine required excavation of tunnels or holes and construction of devices to remove the coal found under the surface. Initial development costs were approximately equal regardless of the amount of coal buried beneath the surface. These fixed costs added an incentive to develop mines reaching large deposits rather than small ones. Variable costs produced more product to sell, but fixed costs did not. Investors preferred large resource supplies.

Time of Use. People avail themselves of energy services in regular cycles, through each day and night period and annually from summer to winter. This regular cycle of periods having high use alternating with periods of low use put a premium on energy sources that could be safely stored until needed. Firewood, coal, oil, and gas lent themselves to easy safe storage, and falling water was reasonably stable except during periods of drought. When uranium joined the energy economy in the 1950s, it, too, was easily stored until needed. Investors knew about these stable, predictable fluctuations, and it was natural to gravitate toward the easily stored primary energy resources. Until the twenty-first century, therefore, time of use was assumed as a criterion supporting investment, but for the most part it was unconscious and invisible.

Practical technologies to use solar radiation and wind did not share this strength in time of use, which made this criterion highly visible.

Geographic Distribution. An energy source found in the company's home country and close to its customers generally has less risky and lower production costs than one located farther away. Energy sources, however, are not evenly distributed across the landscape. Significant oil deposits are found in a few areas, but most areas have little or none. Falling water is plentiful in hilly areas, more diffuse in flat areas. Solar radiation varies strongly by latitude and time of year, even though all parts of the earth have it. Investors recognize the implications and constraints of geographic distribution when making decisions about whether to invest or not.

3.2 Sustainability Criteria

Compared to the long history of the traditional criteria, experience with sustainability began in the 1980s, although path-breaking writings by American and Canadian authors—such as Harrington's *The Other America* [15], Jacobs' *The Death and Life of Great American Cities* [16], Carson's *Silent Spring* [17], Meadows' *The Limits to Growth* [18], and Lovins' *Soft Energy Paths* [19]—paved the way for sustainability as a goal. All of them, plus others from around the world, identified human-created threats to civilization: their economies, the environment, and equity, the pillars united under sustainability in *Our Common Future* [20]. Four criteria identify subsets of attributes included in sustainability.

Protection of Common Resources. Human life and civilization depend upon "common resources" that no person, community, or nation owns and controls. Preservation is a collective responsibility, some at local and some at national or international levels. The most important of these collective responsibilities is the preservation of climate. Similarly, the need to use resources in ways that do not generate conflicts can be achieved only by collective actions. Finally, development and use of energy may threaten health and environment, preventable only by collective actions.

Climate science cannot yet predict the precise rate and magnitude of temperature increases caused by higher CO_2 concentrations in the atmosphere, nor can it precisely predict the consequences and their magnitudes. Nevertheless, the science is mature enough to indicate that potentially catastrophic risks may come from continued emissions of CO_2 and other greenhouse gases. The higher temperatures may melt ice, raise sea levels, alter the frequencies and intensities of floods and droughts, and data from 2016 indicated that human-induced warming, not just natural variability, had caused extreme weather events [21]. Despite lack of precise knowledge about the risks, climate scientists have estimated that global average temperature increases greater than 2 °C are especially dangerous, and the Paris climate agreement embraces a more ambitious goal of limiting increases to 1.5 °C [22].

Similarly, reliance on petroleum and uranium as primary energy resources generates geopolitical tensions that can explode into devastating, international conflicts. Likewise, production and use of all energy sources can damage health and the environment, mostly at local scales [23]. As with climate change, these geopolitical, health, and environmental effects can be limited only by collective actions.

Distribution of Benefits, Costs, and Risks. Sustainability elevates issues of distribution of benefits, costs, and risks as matters that investors must consider. If some individuals or groups enjoy the benefits of energy services but others bear the risks, then inequity and rancor will probably erupt, at least in democratic societies. If all people have voted, leaders will respond to ensure that the distribution of risks and benefits provides acceptable levels of equity among people with different interests and living conditions. Sustainability also identifies intergenerational equity as a significant concern. Future investments, if they are to result in sustainable energy services, must consciously include distribution of benefits, costs, and risks as a criterion.

Ethics. The criterion of distribution suggests measures that are primarily economic and political in nature, but judgments about fairness of distribution inevitably lead to ethical and moral issues. What is the right or ethical pathway for investments? What norms dominate in considerations of morality? Who oversees deciding right from wrong? What rights do those not in charge have if they do not like the judgment call? Questions of ethics and morality regarding the not-yet-born pose especially difficult problems, because these people are not present to speak for their own interests and rights. Sustainable investments in energy must consider ethics, difficult as those considerations may be.

Aesthetics. All energy investments alter the landscape and thus the aesthetic dimensions of the earth. Most energy infrastructure places an industrial object on the ground or in the water, where everybody nearby will see it. People in modern industrial societies are well accustomed to seeing industrial landscapes, but most people do not prefer them. Some such landscapes, in fact, appear ugly to virtually everyone, even though they may be tolerated. Aesthetics raises thorny issues in both the distribution of benefits, costs, and risks and in ethics, further complicating judgments about aesthetics. Sustainable investments in energy must consider aesthetics as a criterion [24].

3.3 Standards

Each cell of the matrix calls for an evaluation of the strength or weakness of an energy source in terms of one of the nine criteria identified above. Investment in that energy source will appear to be a good or strong idea or a bad or weak idea—on that criterion—based on the evaluation. The next section outlines the generation of evaluations and standards in the matrix.

4 Assessment to Determine Standards

4.1 Issues Examined

Three broad areas of inquiry guide the evaluation process. First, investors must know something about the technology involved. Every invention must begin with an idea and imagination. How does it work physically? Is it an idea, but no one has put it into operation? Can anyone imagine that the idea will work? If an investor cannot even imagine that the proposal will work, the purported technology will garner no investments. If it works, what is the stage of the technology's development?

Technologies gain strength as investment targets if they can be imagined or if prototypes demonstrate feasibility. A technology successfully deployed in commerce gains more strength, and a long history of commercial success bears the least risk for the investor. High risks may carry the potential for high rewards, but investors will step forth only when they judge the risks tolerable. As a group, investors will be heterogeneous in their risk tolerances, and qualitative factors will separate them.

Inventions are the key to exploiting an energy source but moving from idea to prototype to deployment to maturation will create industrial concerns with routes to financing development. How large, strong, and stable are the industrial concerns? Are they gaining or losing strength? If so, why? Are they entities of private citizens or governments? Are financial resources to develop the technology adequate? Is the financial basis of the industry voluntary or the result of laws and regulations? How stable, economically and politically, is the financial base of the industry? If government policies undergird the industry, what are the motivations of the policies? Are those motivations stable or labile with changing political situations? Strong industries and financial bases will appear as strengths to investors and encourage investment.

Roles for government policy appear implicitly in considerations of technological inventions and the strengths of industries and financial supports, but policy is so important to the fortunes of investment that it deserves explicit recognition. Investments in energy specifically lie within the core concerns of governments. Consider the following statement by Alberto Cló:

> *Modern states have always considered the energy sector as an exceptional case within the general panorama of industrial activities. Its importance in the development of their economies, in their currency balance, in the distribution of wealth within and among them, and in their status in international relations, was considered so great that states could not be expected to let markets alone achieve goals of general interest that … would not otherwise have been followed spontaneously* [25].

In other words, the role of government policy may play a role as powerful as the physical feasibility of a technology in supporting investment in energy. Investors will see government support as a strength and indifference or hostility as a weakness of a potential investment.

Quantitative measurements for standards offer precision, or sometimes the illusion of precision [26], but subjective elements appear in each of the nine criteria, including the five traditional ones. For this reason, the assessment of each energy source by each criterion must rest on qualitative judgments. Table 2 summarizes the assessment measure made in each cell of the matrix. The symbol, +++, means exceptionally strong, and −−− means exceptionally weak or intolerable. Intermediate assessment scores include ++, +, −, and −−. This assessment method produces ordinal standards that allow ranking of different potential energy sources as targets for investment.

An energy source with mixed properties may be scored with a mixed ranking. For example, an assessment as +/−− indicates that some evidence on the criterion is strong, but other evidence points to a weakness growing weaker.

Because the assessments are qualitative by necessity, different people or groups may make different assessments, reflecting their own situations and interests. Thus, the matrix does not and cannot provide an objective picture of energy investments

Table 2. Symbols and their meanings for the energy–sustainability matrix

Symbol	Meaning
+++	Exceptionally strong
++	Strong, tend to stronger
+	Strong
−	Weak but tolerable
−−	Weak, tend to weaker
−−−	Exceptionally weak, intolerable

commanding universal agreement. Instead, the results of an analysis summarized in the matrix provide a framework for discussions and negotiations among different individuals who will each see the energy investment situation differently.

Assessments in the matrix also change over time. New technology may alter physical feasibility, and changes in economic and political considerations will alter assessments. In addition, changes in risk tolerance or the behavior of energy consumers may alter the outlook for investors. Completion of an analysis produces a labile snapshot, not necessarily a durable image.

4.2 Interpretation of the Matrix

A generic assessment, completed in 2016, compares the strengths and weaknesses of each of the nine primary energy sources plus energy efficiency to provide sustainable energy services for the world. The nine criteria, horizontally at the top, are divided into traditional and sustainability groups. Energy efficiency is placed, in the column on the left, between the nonrenewable and renewable sources. A detailed discussion of each energy source appears elsewhere [27], but the major features of the completed matrix become visible by dividing the matrix into four quadrants (Table 3).

Table 3. The energy–sustainability matrix, with standards reflecting one person's assessments

CRITERIA	Profits and Risks	Security of Access and Profits	Size of Resource	Time of Use	Geographic Distribution	Protection of Common Resources	Distribution of Benefits, Costs, and Risks	Ethics	Aesthetics
SOURCE	TRADITIONAL CRITERIA					SUSTAINABILITY CRITERIA			
Coal	+++/--	+++	+++	+	++/-	---	+/---	---	---
Petroleum	+++/--	+/--	+	+++	---	---	+/---	---	-/--
Gas	+++/-	+/--	+	+++	--	---	---	---	-/--
Uranium	+/---	+++	+	+	-	+/--	+/---	---	-
Efficiency	+++	+++	+++	+++	+++	+++	+/-	+++	+++
Solar	++/-	+++	++	-	++/-	+++	+++	+++	+/-
Wind	++/-	+++	++	-	++/-	+++	+++	+++	+/-
Hydropower	+/-	+++	++	+++	+++	+++/-	+++/-	+++/-	+/-
Biomass	+/--	+++	--	+++	+++	+/--	+/--	+/--	+
Geothermal	+/-	+++	+/+++	+++	+++	+++	+++/-	+++	-

Nonrenewable Sources, Traditional Criteria. The upper left quadrant displays the assessment of the sources currently dominating the world's energy economy, which has developed over the past 300 years. Most of the cells have at least one +, and many have three (+++). Some negative assessments also appear, reflecting various tensions surrounding the uses of each energy source. For example, investments in coal in some cases are very strong (+++), but its falling sales due to competition with natural gas and renewable energy sources have led to bankruptcies and concerns for investors (−−). Petroleum, the largest single source of energy and the current bedrock of mobility, also remains a lucrative investment in some instances (+++), but it suffers growing competition due to electric vehicles (−−) and its limited geographic distribution remains a source of geopolitical tensions (−−−).

Nonrenewable Sources, Sustainability Criteria. In the upper right quadrant, the nonrenewable sources do not show much strength. The three fossil fuels, for example, each show a powerful weakness (−−−) on protection of common resources, because it is currently impossible to use most of them without emitting CO_2 to the atmosphere. Similarly, they show predominantly weak scores (−−−) on distribution of costs, risks, and benefits, because most of the uses benefit people in the rich countries, but all people—even if they do not use much fossil fuel—will endure the risks of climate change. Uranium, in contrast, shows a strength (+) due to its role in generating electricity with low emissions of CO_2, but the dangers of catastrophic accidents, inability to politically resolve waste disposal issues, and the intrinsic production of materials for the proliferation of nuclear weapons show great weakness (−−−).

Renewable Sources, Traditional Criteria. In the lower left quadrant, the renewable sources show a mixed situation, with strengths intermingling with weaknesses. The nonrenewable sources tend to rank higher on profits and risks, the major criterion guiding investments for the last 300 years. Renewable sources are approaching economic competitiveness, and thus attractiveness for investment, but it is not yet clear they can attract investment without subsidies and policy directives. Moreover, solar radiation and wind still have constraints on time of use because of intermittency. All the other nonrenewable sources have a +++ rank, matching the strengths of petroleum and gas. Overall, most investors might see the nonrenewable sources as better investment targets than renewable sources, but changes are clearly in progress.

Renewable Sources, Sustainability Criteria. In the lower right quadrant, the renewable sources show much greater strength. They also show, compared to the nonrenewable energy sources, a tremendous contrast, with many +++ rankings. Some also exhibit mixed scores. Hydropower, for example, scores highly on each of the sustainability criteria, but it also has some weaknesses. It scores +++ on protection of common resources, due to low CO_2 emissions, but it also has a weakness (−) due to its disruption of rivers as a common resource. Controversies over the destruction of fisheries in the Pacific Northwest of the United States by dams, for example, have shown that, for some people, hydropower is far from a benign energy source [28].

Energy Efficiency. Efficiency is uniquely strong as a target for investment on both traditional and sustainability criteria, far higher than any of the primary energy sources. These scores underline the common wisdom to invest in efficiency first, last, and always.

4.3 Decision-Making Informed by the Matrix

The matrix is a device to organize discussions among people about energy, and it focuses attention on the properties of primary energy sources. Its utility stems from the analytical framework identifying investment as the key decision in the development and use of primary energy sources. These decisions involve difficult dilemmas and contradictions, even if the decision-makers include only people from industry and finance. Governments, however, have always participated in decision-making, and incorporation of criteria for sustainability broadens the legitimate players to the entire public. Inevitable conflicts arise between competing stakeholders, who may find reasonable and factual evidence supporting different weights for the various criteria and multiple outcomes for assessments of competing investment opportunities.

This short paper cannot delve into the intricacies of decision-making involving contradictory assessments of the multiple goals specified by the criteria. Suffice it to say that extensive literature exists on key topics such as risk assessment [29] and the need for effective, broad public participation in decision-making [30]. Utility of the matrix for building sustainable energy systems will depend upon skillful use of what we know about these topics.

5 Conclusions

Decisions about investments are key in building, maintaining, and changing the energy on which fully modern civilization depends. Without investments, people cannot enjoy energy services, and investors must choose among nine primary energy sources, plus energy efficiency, to make these services available. Decisions will always rest on multiple considerations, some of which have precise, quantitative measures. For example, the number of dollars required and the question of technical functionality can usually be known with some precision.

Other factors, however, will always remain subjective, subject to different assessments by equally rational and competent individuals. These factors, including political issues, risk estimations, ethics, and aesthetics, will frequently be assessed differently. Moreover, the nine criteria in the energy–sustainability matrix may carry different comparative weights by different individuals. Some, for example, will weigh the criterion of profits more heavily, and others will see protection of common resources or ethics as more important.

Due to subjective factors, decision-making about investments in future energy economies is a sociopolitical and moral arena in which quantitative engineering and scientific concerns also play a fundamental role. The matrix identifies key issues and invites different individuals to make assessments from their own perspectives. It serves as a framework for discussions, and negotiations in case different people have different attitudes about prospective investments and proposed energy reforms. The matrix does not settle disputes, but it frames debates in systematic ways that help people find common ground on one of the most important issues facing the world in the twenty-first century.

Acknowledgments. I am indebted to Dr. Tom Womeldorff for his advice on improving the specificity of the word "sustainability." I am grateful to four anonymous reviewers who provided comments that improved the narrative. Thanks also go to the organizers of the 2018 Energy and Sustainability Conference for including this paper in the program.

References

1. Fell, M.J.: Energy services: a conceptual review. Energy Res. Soc. Sci. **27**, 1129–1140 (2017)
2. International Energy Agency: Key World Energy Statistics, p. 8. International Energy Agency, Paris (2017). (percent calculation by author)
3. Day, R., Walker, G., Simcock, N.: Conceptualising energy use and energy poverty using a capabilities framework. Energy Policy **93**, 255–264 (2016)
4. Sterling, A.: Transforming power: social science and the politics of energy choices. Energy Res. Soc. Sci. **1**, 83–95 (2014)
5. McCollum, D.L., et al.: Connecting the sustainable development goals by their energy interlinkages. Environ. Res. Lett. **13**, 033006 (2018). https://doi.org/10.1088/1748-9326/aaafe3
6. Wrangham, R., Carmody, R.: Human adaptation to the control of fire. Evol. Anthropol. **19**, 187–199 (2010)
7. Perkins, J.H.: Changing Energy: The Transition to a Sustainable Future, pp. 24–49. University of California Press, Oakland (2017)
8. Perkins: Changing Energy, pp. 136–165 (2017)
9. Intergovernmental Panel on Climate Change: Climate Change 2013: The Physical Science Basis, Cambridge University Press, New York, pp. 676–678, 697 (2013)
10. International Institute for Sustainable Development: What Is Sustainable Development? 23 March 2018. www.iisd.org/sd#one
11. Womeldorff, T.: Personal Communication, 9 April 2018
12. Perkins: Changing Energy, pp. 202–209 (2017)
13. Wrigley, E.A.: Energy and the English Industrial Revolution. Cambridge University Press, Cambridge (2010). 272 pp
14. Perkins: Changing Energy, p. 138 (2017)
15. Harrington, M.: The Other America: Poverty in the United States. Macmillan, New York (1961). 191 pp
16. Jacobs, J.: The Death and Life of Great American Cities. Vintage Books, New York (1961). 458 pp
17. Carson, R.: Silent Spring. Houghton Mifflin Co., Boston (1962). 368 pp
18. Meadows, D., et al.: The Limits to Growth. Universe Books, New York (1972). 205 pp
19. Lovins, A.: Soft Energy Paths: Towards a Durable Peace. Harper & Row, New York (1979). 231 pp
20. World Commission on Environment and Development: Our Common Future. Oxford University Press, Oxford (1987). 383 pp
21. Herring, S.C., Christidis, N., Hoell, A., Kossin, J.P., Schreck III, C.J., Stott, P.A.: Introduction to explaining extreme events of 2016 from a climate perspective. Bull. Am. Meteorol. Soc. **99**, S1–S6 (2018)
22. United Nations: Paris Agreement (2015). Article 2
23. Perkins: Changing Energy, pp. 137–155 (2017)
24. Apostol, D., et al.: The Renewable Energy Landscape: Preserving Scenic Values on Our Sustainable Future. Routledge, Abingdon (2017). 286 pp

25. Cló, A.: Oil Economics and Policy, p. 40. Kluwer Academic, Dordrecht (2000). English translation of Economía e politica del petrolio sponsored by the European Society for Scientific Publications
26. Porter, T.: Trust in Numbers: The Pursuit of Objectivity in Science and Public Life. Princeton University Press, Princeton (1995). 310 pp
27. Perkins: Changing Energy, pp. 219–268 (2017)
28. Mulier, V.: Recognizing the full scope of the right to take fish under the Stevens Treaties: the history of fishing rights litigation in the Pacific Northwest. Am. Indian Law Rev. **31**, 42–92 (2006/2007)
29. Stern, P.C., Fineberg, H.V. (eds.): Understanding Risk: Informing Decisions in a Democratic Society. National Academy Press, Washington (1996). 249 pp
30. Dietz, T., Stern, P.C.: Public Participation in Environmental Assessment and Decision Making. The National Academies Press, Washington (2008). 305 pp

Author Index

A. Vasel and D. S.-K. Ting (Eds.): EAS 2018, SPE, pp. 183–183, 2019.
https://doi.org/10.1007/978-3-030-00105-6