

$$W = P_n \cdot k_n \cdot T_{м.ф}, \quad (4)$$

где P_n – номинальная мощность двигателя насоса, кВт; k_n – коэффициент использования, $T_{м.ф}$ – месячный фонд рабочего времени электрооборудования, ч.

На рис. 2 демонстрируются результаты прогнозирования электропотребления насосных агрегатов первого подъема воды в зависимости от исследуемого года.



Рис. 2. Сводный график зависимости потребляемой мощности от исследуемого года

По приведенным графикам теперь можно сделать вывод о том, что в период с 2019 по 2023 г. режимы работы для водозабора являлись стабильными. Однако, начиная с января 2024 г. и до настоящего времени, возникшее аварийное событие привело к значительному снижению энергетических показателей и, как следствие, уменьшению электропотребления. Полученная информация может пригодиться для эксплуатационного персонала водозабора в части анализа влияния загрузки оборудования на энергетическую эффективность, что особенно актуально при подборе стратегий управления насосными станциями. Благодаря представленным результатам, появляется возможность прогнозировать режимы работы оборудования. В связи с чем возникает задача проведения дальнейших исследований по выявлению закономерностей влияния загрузки водозабора на показатели, характеризующие графики электрических нагрузок.

Литература

1. Optimize the cost of paying for electricity in the water supply system by using accumulating tanks / A. Kapanski, N. Hruntovich, S. Bakhur [et al.] // E3S Web of Conferences, Prague, 14–15 May 2020. – Prague, 2020. – P. 01065. – DOI 10.1051/e3sconf/202017801065. – EDN CSSMKC.

МЕТОДИКА ПРОВЕДЕНИЯ ОДНОФАКТОРНОГО ДИСПЕРСИОННОГО АНАЛИЗА ДЛЯ ВЫЯВЛЕНИЯ ГРУПП ПОТРЕБИТЕЛЕЙ ЧАСТНОГО ЖИЛОГО СЕКТОРА

В. В. Павлов, Д. И. Веремеева

Учреждение образования «Гомельский государственный технический
университет имени П. О. Сухого», Республика Беларусь

Научный руководитель А. А. Капанский

Представлен алгоритм проведения однофакторного дисперсионного анализа для данных об электропотреблении частного жилого сектора, который в дальнейшем позволит

найти группы, имеющие значительное сходство, для формирования впоследствии их типовых профилей.

Ключевые слова: электропотребление, бытовой частный дом, однофакторный дисперсионный анализ, множественные сравнения.

Цель исследования заключается в выявлении характера электропотребления для последующего определения пропускной способности сети ввиду роста установленной мощности потребителей. Данные для проведения исследования были получены в ходе выполнения на кафедре «Электроснабжение» НИР «Разработка методики оценки пропускной способности электрических сетей 0,4 кВ в условиях роста электропотребления бытовых потребителей частного жилого фонда». Общий объем данных по электропотреблению одной из деревень составил 18 144 строк времени и 41 столбец домов. Период снятия данных – с августа 2023 г. по август 2024 г.

Очистка массива данных от нехарактерных потребителей. Годовое электропотребление каждого дома представлено на гистограмме (рис. 1). Видно, что некоторые дома имеют незначительное потребление, которое для репрезентативности выборки необходимо исключить из анализа:

1. Минимальное и практически нулевое электропотребление имеют дома 27, 52, 54 и 99, которые исключаются из анализа.

2. В ходе рассмотрения всего электропотребления в выборке было решено установить границу годового электропотребления на 150 кВт/год или 75,0 кВт·ч/год. Выбор граничного условия обусловлен нагрузкой дома 69, которая имеет ярко выраженный характер в летний период, а в зимний период отсутствует. На основании этого из анализа также исключается дом 107.

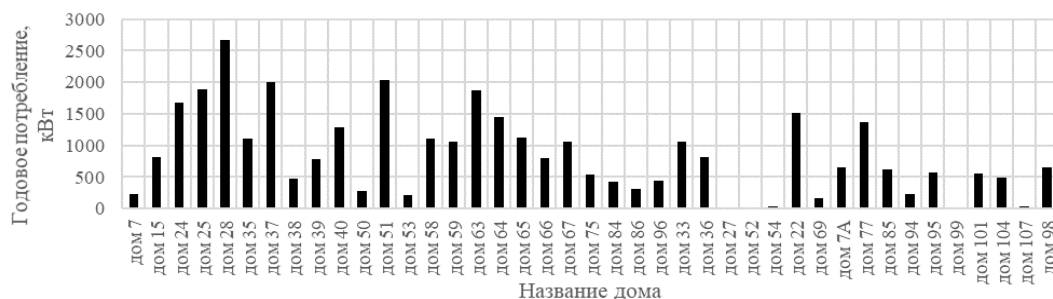


Рис. 1. Анализ годового электропотребления домов

Подготовка данных к проведению ANOVA (однофакторного дисперсионного анализа). Рассмотрение результатов ANOVA. Проведение однофакторного дисперсионного анализа (в дальнейшем – ANOVA) позволяет определить влияет ли фактор электропотребления на различие групп между собой. Данный тест дает возможность определить дальнейшую стратегию формирования типовых профилей.

Перед проведением ANOVA необходимо проверить данные на соответствие критериям, диктуемых данным тестом:

- 1) данные должны иметь количественный характер и должны быть непрерывны;
- 2) выборки должны быть независимы;
- 3) наблюдения должны быть независимы;
- 4) выборки должны соответствовать нормальному распределению;
- 5) должно наблюдаться равенство дисперсий.

Проверка данных на количественный характер и их непрерывность. Данные имеют количественный характер, но непрерывность данных не наблюдается из-за наличия пустот в массиве данных, которые могут быть связаны с ошибками в работе АСКУЭ. Для их заполнения предлагается разбить годовой объем данных по месяцам, а на их местах использовать средние значения, основанные на часе электропотребления и типу дня недели (выходной или будний), усредненные на интервале месяца. Для выполнения этой задачи на языке программирования VBA был реализован код, выполняющий разбивку данных по месяцам и заполняющий пустоты в каждом месяце. Параллельно с выполнением алгоритма формируется таблица, в которой собирается информация об удельном количестве пустых значений в выборке конкретного дома в определенный месяц. Программа также выбирает оптимальный месяц для проведения ANOVA и выделяет этот месяц цветом, при этом данные выделяются по принципу: зеленый – минимальный, желтый – медианный, красный – максимальный. Далее характерный месяц очищается от домов с большим количеством удельного веса нулевых и пустых значений (принята граница 15 %). Так, из ANOVA было решено исключить дома: 7, 39, 58, 75, 84, 36, 69, 7А, 94, 95 (см. таблицу).

Таблица результатов обработки массива

Месяц	Номер дома																																				
	дом 7	дом 15	дом 24	дом 25	дом 28	дом 35	дом 37	дом 38	дом 39	дом 40	дом 50	дом 51	дом 53	дом 58	дом 59	дом 63	дом 64	дом 65	дом 66	дом 67	дом 75	дом 84	дом 86	дом 96	дом 33	дом 36	дом 22	дом 69	дом 7А	дом 77	дом 85	дом 94	дом 95	дом 101	дом 104	дом 98	
январь	80	77	77	77	77	77	77	85	79	77	100	77	100	77	77	77	77	77	89	77	82	94	94	83	77	80	78	100	77	77	100	97	87	83	79	77	
февраль	3	0	0	0	0	0	0	33	29	0	99	0	99	0	1	0	0	0	49	1	16	38	39	35	0	24	2	99	0	0	83	59	25	17	26	0	
март	1	0	0	0	0	0	0	34	22	0	99	0	99	1	0	0	0	0	57	0	17	33	25	11	0	24	1	99	0	0	39	60	27	8	28	0	
апрель	12	0	0	1	0	0	0	19	24	0	93	0	96	0	1	0	0	0	59	0	17	1	8	12	1	12	0	72	0	0	9	23	29	6	32	1	
май	87	84	84	84	84	84	84	86	87	84	87	84	99	84	84	84	84	84	94	84	88	84	84	84	90	84	86	84	100	84	84	85	91	89	85	88	84
июнь	64	4	1	1	0	1	1	9	19	0	0	0	47	5	1	0	1	0	64	0	16	22	2	2	0	19	0	92	15	0	2	19	26	8	6	1	
июль	85	13	14	17	14	16	15	16	27	14	13	67	99	17	20	14	15	16	79	13	21	19	17	17	16	23	14	22	14	14	18	27	33	18	18	14	
август	73	1	0	0	0	0	0	3	12	0	1	42	61	4	42	0	3	0	66	3	19	7	1	30	4	13	0	1	41	0	28	48	26	7	5	0	
сентябрь	40	0	0	0	0	0	0	17	5	1	35	0	13	0	100	0	0	0	59	0	19	3	3	47	8	20	0	0	93	0	60	51	30	17	10	0	
октябрь	15	0	0	0	0	0	0	19	7	0	88	0	18	0	100	0	0	0	53	0	20	4	12	24	4	18	0	11	89	0	84	44	28	12	25	0	
ноябрь	77	74	74	74	74	74	74	79	76	74	100	74	80	76	100	74	76	74	88	74	83	84	78	83	74	77	74	77	91	74	100	95	82	77	84	78	
декабрь	6	0	0	0	0	0	0	31	8	0	98	0	94	0	0	0	0	0	51	0	18	89	47	7	0	13	0	58	4	0	100	90	31	14	1	0	

Проверка данных выборки на независимость, а также условия независимости наблюдения. Потребление электроэнергии одного дома никак не влияет на потребление другого, что указывает на независимость данных. Наблюдения также являются независимыми, так как на каждого потребителя установлен отдельный прибор учета (электросчетчик).

Проверка данных выборки на соответствие нормальному распределению. Проверку выборок на нормальность можно провести разными способами, но для быстрого и наглядного анализа данных был выбран графический метод – с помощью графиков «ящик с усами» (рис. 2, а). Видно, что данные имеют некоторое количество выбросов у каждого дома, которые находятся за границей 1,5 межквартильного размаха. Для их исключения в VBA был реализован алгоритм, заменяющий значения выбросов на значения линейной интерполяции между предыдущим и следующим получасовым потреблением, не выходящим за межквартильный размах (рис. 2, б).

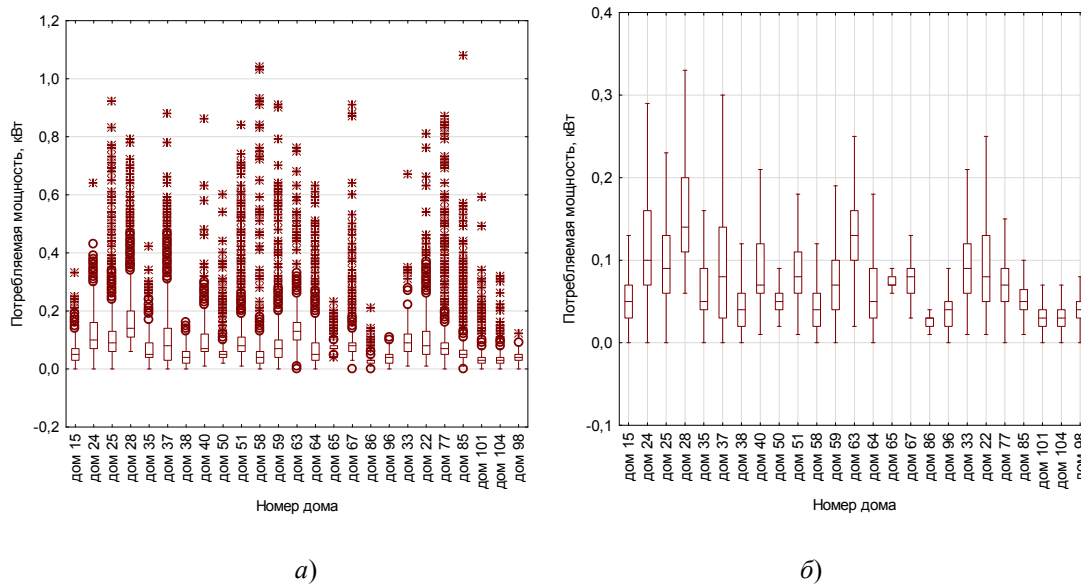


Рис. 2. Диаграмма «ящик с усами» до (а) и после исключения выбросов (б)

Как видно на диаграмме рис. 2, многие данные имеют ненормальное распределение. Так, дом 24 имеет значительно скошенную медиану потребления от середины и неравномерную длину «усов», что отклоняет теорию о нормальности. Тогда для выполнения ANOVA предлагается использовать его непараметрический аналог, а именно – тест Краскела–Уоллиса, основными условиями которого являются все условия ANOVA, кроме условий, подразумевающих нормальность данных.

Формирование гипотез ANOVA и определение уровня значимости. Нулевая гипотеза или H_0 : заключается в том, что статистически значимые отличия между группами (т. е. домами) по фактору (электропотребление) отсутствуют, и сам по себе дом никак не влияет на электропотребление.

Конкурирующая гипотеза или H_1 : заключается в том, что существуют статистически значимые отличия между группами (т. е. домами) по фактору (электропотребления), и дом значительно влияет на электропотребление. Уровень значимости для большинства технических работ принимается на уровне 5 %.

Проведение ANOVA. Для проведения теста Краскела–Уоллиса предлагается использования программного комплекса Statistica, в котором можно легко получить данные теста. Итогом теста является таблица с выведенным уровнем значимости, в нашем случае оно имело значение $p = 0,000$, что меньше принятого нами значения 0,05. Это означает, что группа (дома) оказывает существенное влияние на фактор (электропотребление).

Литература

1. Cohen, A. C. Truncated and censored samples: theory and applications / A. C. Cohen. – CRC press, 2016. – 313 p.
2. Ивановский, Р. И. Теория вероятностей и математическая статистика. Основы, прикладные аспекты с примерами и задачами в среде Mathcad / Р. И. Ивановский. – СПб. : БХВ-Петербург, 2008. – 528 с.