



Марин Юрий Алексеевич — кандидат технических наук, старший научный сотрудник (доцент), в настоящее время работает на кафедре «Инженерная геодезия», участвует в работе НИЛ по борьбе с заносами и лавинами на железных и автомобильных дорогах. В 1973 г. защитил кандидатскую диссертацию на тему «Исследование методов защиты лавиноопасных участков ж.-д. пути на Сахалине». Автор около ста работ, в том числе двух учебных пособий для курсового и дипломного проектирования, десяти методических разработок и рекомендаций.

E-mail: marin5635@yandex.ru

УДК 625.551.578.4.322

Ю.А. МАРИН

КОРРЕЛЯЦИОННО-РЕГРЕССИОННЫЙ АНАЛИЗ ВЗАИМОСВЯЗИ ПЛОТНОСТИ И СОПРОТИВЛЕНИЯ СДВИГУ В СФОРМИРОВАВШЕЙСЯ ТОЛЩЕ СНЕЖНОГО ПОКРОВА

Приведены результаты статистического анализа взаимосвязи физико-механических свойств снега — сопротивления сдвигу и плотности. Эти параметры измерены попарно в 10-сантиметровых слоях снежного покрова в течение 13 зим с декабря по апрель в горной лавиноопасной местности Южного Сахалина. Получены кривые распределения величин плотности, измеренных для каждого из 12 значений сопротивления сдвигу в диапазоне от 0,49 до 1,67 кПа. Выполнена статистическая оценка коэффициентов корреляции. Представлены статистические характеристики коэффициентов корреляции. С помощью дисперсионного анализа сопоставлены зависимости линейного и параболического характера. Проверена гипотеза о принятом виде функциональной зависимости. Изложена схема возможной реализации подобного исследования, приведены выводы по результатам анализа.

Ключевые слова: *физико-механические свойства снега, снежные лавины, сопротивление сдвигу, плотность, статистический анализ, коэффициент корреляции.*

Существует достаточное количество теоретических и экспериментальных свидетельств, показывающих качественную связь между такими параметрами снега, как плотность и сопротивление сдвигу. Доказано, что эти параметры в значительной степени зависят от времени и температуры. Плотность влияет на нормальное напряжение к поверхности сдвига. Полевые измерения сопротивления снега сдвигу с помощью сдвиговой рамки выполняются для оценки устойчивости снежной доски на склоне и возможного выявления участков возрастания напряжений в периферийной части снежного пласта. Рост этих напряжений способен вызвать подвижки снега и сход снежной лавины.

Сопоставление данных измерений плотности и сопротивления сдвигу в одном и том же горизонте снежной толщи показывает, что каждому значению плотности ρ соответствует неопределенное количество значений сопротивления сдвигу S . Средние из этих значений одной переменной зависят от значений другой. Можно считать, что существует почти функциональная зависимость средних значений сопротивления сдвигу от соответствующих значений плотности и наоборот. В этом случае можно утверждать, что между рассматриваемыми параметрами снега существует корреляционная зависимость.

Коэффициент корреляции R , определяемый как число, характеризующее степень связи между двумя переменными, зависит от количества пар наблюдений или измерений и принятого уровня значимости P (0,01, 0,025 или 0,05) и его величина находится

в диапазоне $-1 < R < 1$. Коэффициент корреляции не зависит от размерности исследуемых параметров и начала отсчета.

При изучении особенностей корреляционной связи функциональной зависимости $C = f(p)$ правомерно рассматривать зависимость $p = f(C)$. Для этого исходные данные аргумента (например p) дополняют его средними значениями, а исходные данные аргумента C — его средним значением $\bar{C}$. Для обработки получают две таблицы данных $(p, \bar{C})$ и $(C, \bar{p})$, которые определяют эмпирическую регрессию C по p и эмпирическую регрессию p по C .

Корреляционно-регрессионный анализ позволяет более точно определить вид или тип связи между двумя рядами переменных. Картина этой связи между параметрами или свойствами, присущими какому-либо телу или среде, часто носит не четко детерминированный, а стохастический (вероятностный) характер [1].

Исходя из практических удобств при большом количестве переменных в ряду (более 100) и учитывая активное влияние слабо контролируемых факторов, весь диапазон измерений делят на классы, формируя выборку, и по ней ведут расчеты [2, 3].

Построение зависимости между двумя рядами переменных может быть реализовано по различным схемам [1]. По схеме 1-й зависимость между неслучайными переменными не требует использования вероятностно-статистических теорий. По схеме 2-й зависимость случайной переменной η от неслучайной переменной x сопровождается рядом неконтролируемых факторов и поэтому при каждом фиксированном значении x соответствующее значение зависимой переменной подвержено некоторому случайному разбросу. Обычно полагают, что η может быть разложена на сумму двух слагаемых. Первое — неслучайная часть — определяется некоторой линейной функцией от x , а второе — случайная часть — учитывает случайный характер η . Согласно 3-й схеме исследуемые величины зависят от совокупности неконтролируемых факторов и по своей физической сущности являются случайными. По условию схемы 4-й исследуемые величины x и y не случайны, но могут быть измерены только с некоторыми случайными ошибками измерения δ_x и δ_y . В этом случае приходится иметь дело с выборочными значениями случайных величин $\varepsilon = x + \delta_x$ и $\varepsilon = y + \delta_y$.

Предварительные представления о корреляционных связях между плотностью и сопротивлением сдвигу для различных типов снега и их модульных значений приведены в [4], где значение плотности принято за независимую переменную, а сопротивление сдвигу — за функцию. Оценивая полученные зависимости с позиции профессионально-технического и историко-генетического анализа, следует учесть, что в процессе формирования снежной толщи она сама и ее физико-механические свойства подвергаются влиянию температурных и теплофизических факторов, воздействию внешних нагрузок. Свойства снега зависят от совокупности неконтролируемых или в слабой степени контролируемых факторов и по своей физической сущности являются случайными величинами. Анализ таблиц $C = f(p)$ и $p = f(C)$ показывает, что когда величина p растет, то отдельные значения C могут и расти и убывать, одному и тому же значению p могут соответствовать как большие, так и малые значения C . То же самое имеет место для зависимости $p = a(C)$. В то же время средние значения одной величины, соответствующие значениям другой, оказываются в известной зависимости от второй величины.

В теории вероятностей применяется закон, согласно которому значения какой-то величины при достаточно большом числе измерений с некоторой относительной точностью группируются около наиболее вероятного среднего значения и который находит широкое применение в физике. Такая величина названа случайной и является

переменной, значения которой зависят от случая, для нее определена функция распределения вероятностей [5, 6].

Полученные корреляционные связи между плотностью и сопротивлением сдвигу для разных типов стратиграфического строения снега и разных временных периодов его залегания (от 1 до 5 мес.), приведенные в работе [4], свидетельствуют о том, что эти связи являются случайными. Они характеризуют соотношение схожих по своей физической сущности двух рядов переменных, подвергаемых влиянию совокупности неконтролируемых или слабо контролируемых факторов. К ним относятся:

- 1) время формирования и роста высоты снежного покрова;
- 2) условия залегания снежного покрова (горизонтальная, наклонная поверхность и ее крутизна, наличие неподвижного препятствия перед массой снега и др.);
- 3) изменение температурного градиента в снежном покрове под воздействием температуры воздуха, интенсивности и величины выпадения твердых осадков, метелевого переноса снега, теплового потока от подстилающей поверхности и др.;
- 4) интенсивность процессов метаморфизма;
- 5) воздействие внешних и внутренних гравитационных нагрузок;
- 6) ошибки, допущенные при измерениях C и p .

Случайные изменения перечисленных факторов, из которых определяющим, по мнению многих исследователей, являются процессы конструктивного метаморфизма, приводят к случайным колебаниям величин обоих рядов переменных. Но расположение точек в поле распределения оказывается вытянутым в направлении, не параллельном ни одной из координатных осей. Функция, с помощью которой можно описать характер этого направления, а значит и взаимосвязь между исследуемыми переменными при признании их случайными величинами, должна быть линейной относительно оцениваемых параметров. По утверждению специалистов, применение метода наименьших квадратов, с помощью которого оценивают параметры, состоящие в определенной функциональной зависимости, может быть реализовано и для случая, когда переменные находятся в нелинейной зависимости (степенной, параболической, логарифмической и т.п.). Но для описания криволинейных зависимостей не рекомендуется использовать параболические зависимости высоких порядков, поскольку это значительно увеличивает объем вычислений и при ограниченном экспериментальном материале заметно снижает точность эмпирической регрессионной кривой [1, 2].

В статье рассмотрены результаты статистической обработки значений сопротивления сдвигу C и плотности p , измеренных в 10-сантиметровых слоях снежной толщи шурфов, заложенных с 1 декабря по 31 марта в течение 13 зим. Построены эмпирические кривые регрессии и вычислены коэффициенты корреляции функций $C = f(p)$ и $p = f(C)$ для трех диапазонов высоты снежного покрова — 50–100 см, 110–140 см и 150–350 см. При первичной статистической обработке измерений рекомендуется исключать резко выделяющиеся результаты. Детальное рассмотрение полей рассеяния диапазона значений сопротивления сдвигу 0,49–1,67 кПа и соответствующих им плотностей показало, что одному из значений C может соответствовать до 40 значений плотности величиной от 0,05 до 0,315 г/см³ [4, табл. 3]. В расчеты введено понятие частоты случайной величины, как отношения ее повторений к общему числу измерений. Частость не тождественна вероятности и может меняться в зависимости от условий проведения различных серий измерений.

Получено 12 кривых с точностью подбора интервала выборки около 5 %, которые описываются асимметричной параболической кривой 2-го порядка (рис. 1–6). Диапазон изменения коэффициентов корреляции от 0,75 до 0,94.

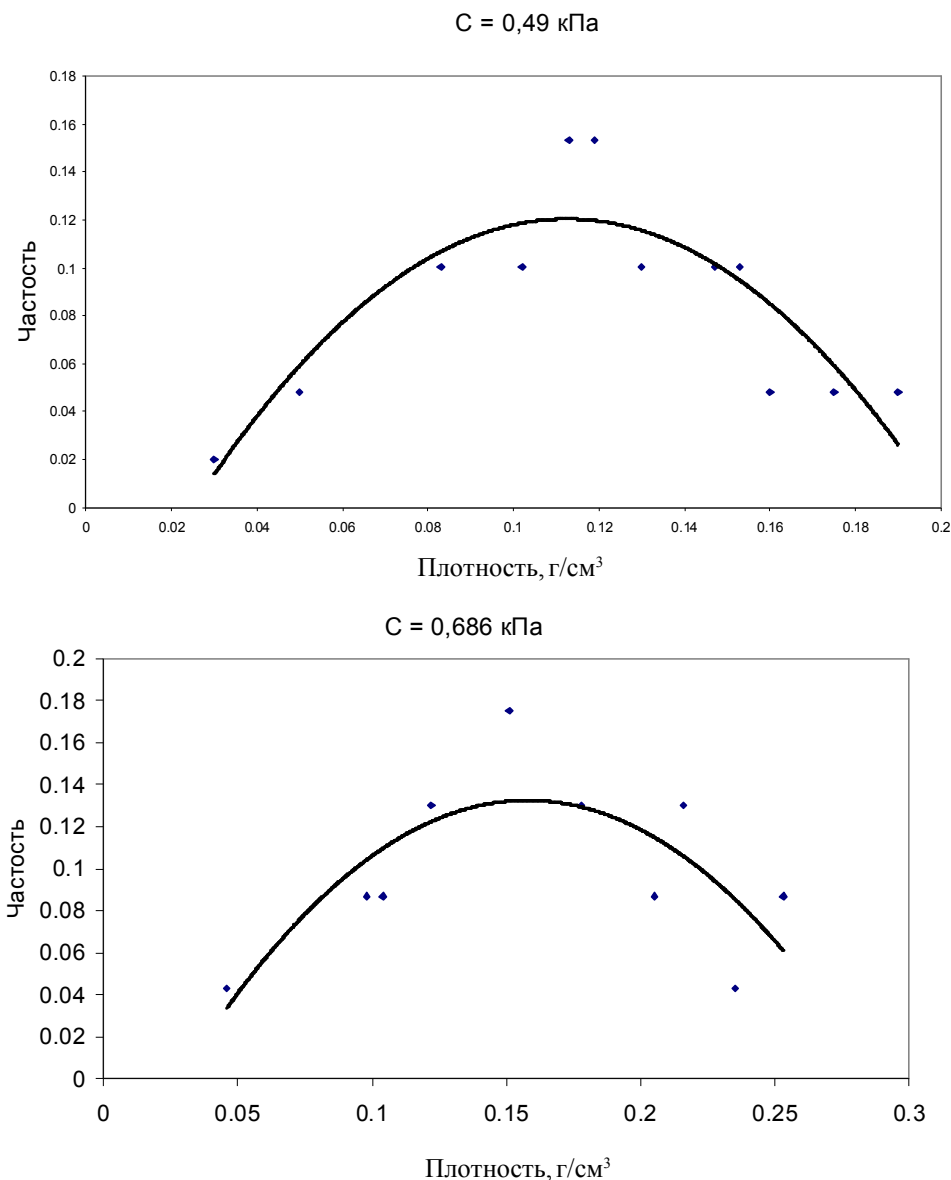


Рис. 1. Распределение величины плотности при сопротивлении сдвигу 0,49 и 0,686 кПа

Анализ теоретических кривых распределения частоты (см. рис. 1–6) показал, что с ростом сопротивления сдвигу, измеренного сдвиговой рамкой, растет диапазон значений плотности снега. Максимальному теоретическому значению частоты сопротивления сдвигу C соответствует некоторое среднее значение плотности p . Взаимосвязь между C и p в измеренных диапазонах представлена зависимостью (1) и показана на рис. 7.

$$p = -0,0676C^2 + 0,2315C + 0,0186. \quad (1)$$

Коэффициент корреляции $R = 0,9682$.

При увеличении количества измерений p возможно приближение формы кривых к более сложной с уравнением вида:

$$y = y_0 x^{h_1} (x-1)^{h_2}, \quad (2)$$

где $h_2 > 0$, а остальные параметры вычисляют по специальным формулам. Такую кривую относят к кривым Пирсона У1 типа [3]. На рис. 8 показана кривая, вид которой соответствует выражению (2).

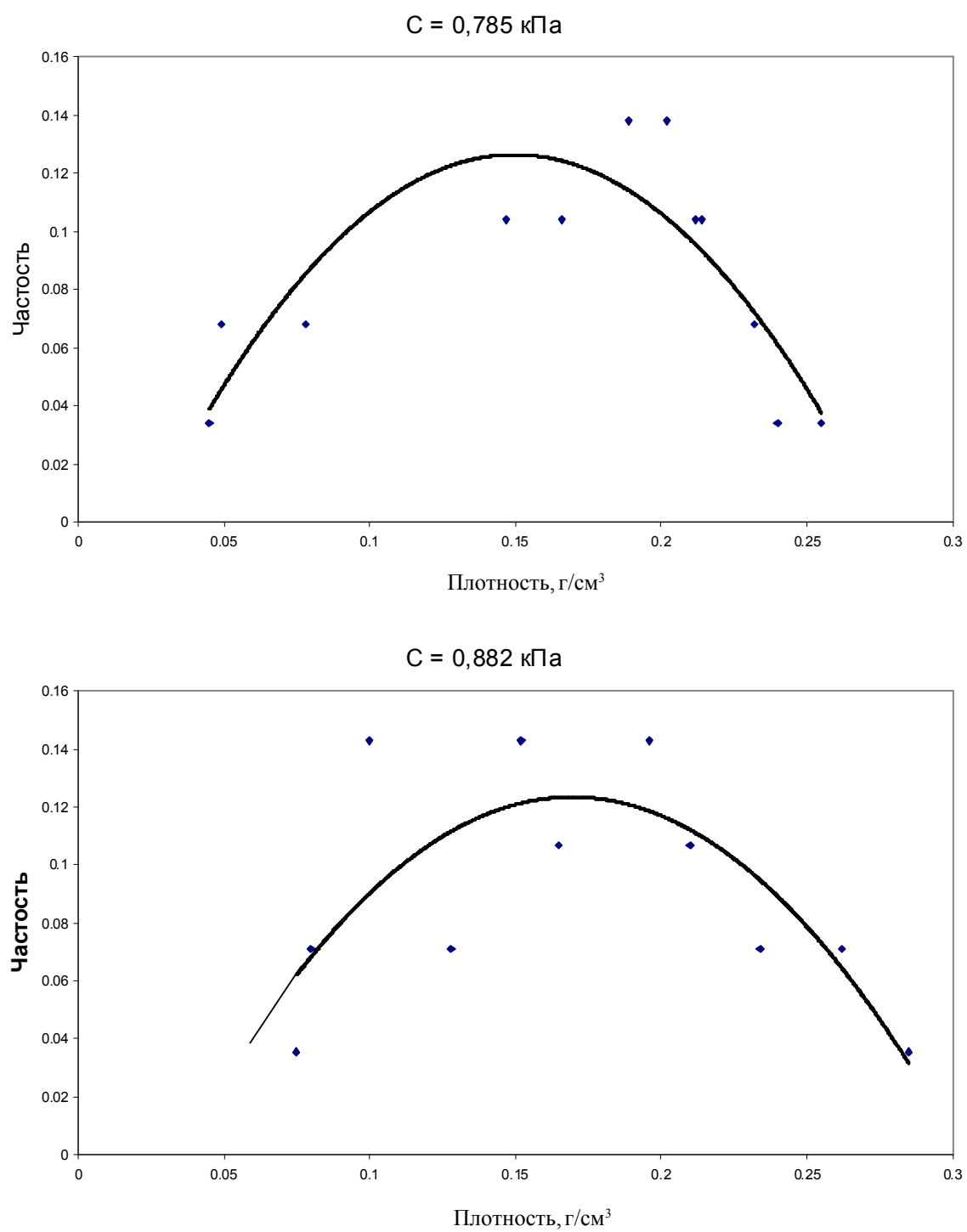


Рис. 2. Распределение величины плотности при сопротивлении сдвигу 0,785 и 0,882 кПа

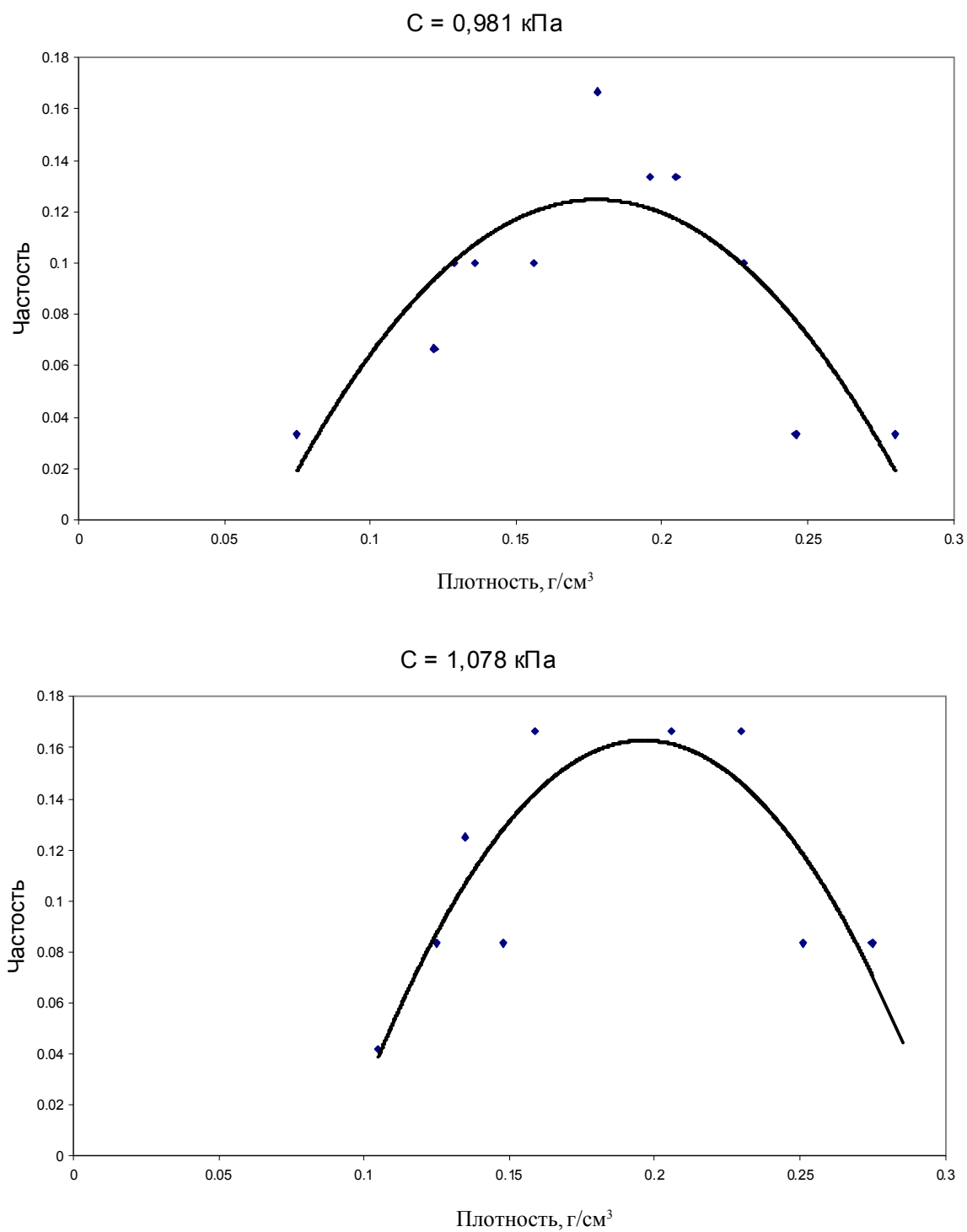


Рис. 3. Распределение величины плотности при сопротивлении сдвигу 0,981 и 1,078 кПа

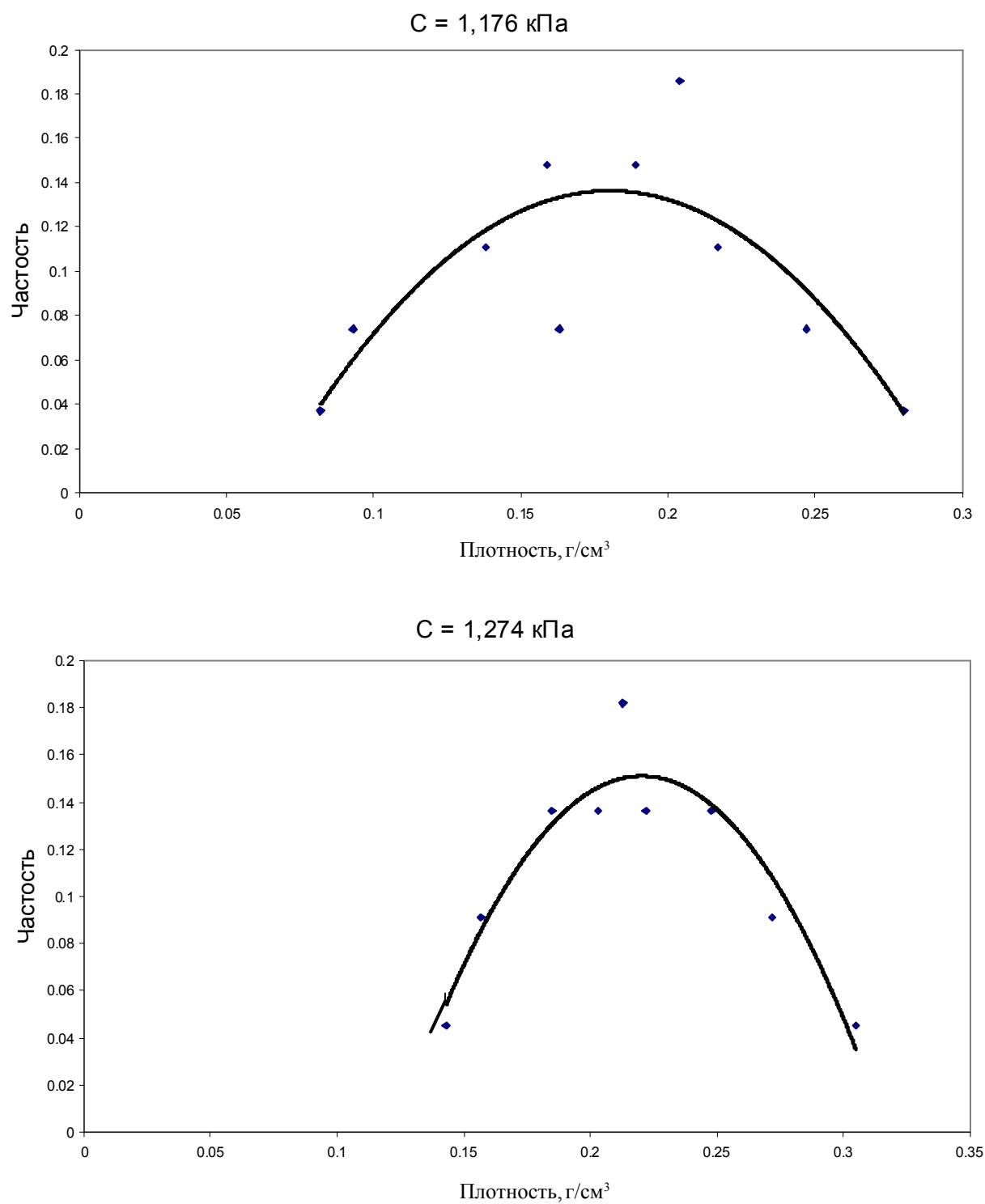


Рис. 4. Распределение величины плотности при сопротивлении сдвигу 1,176 и 1,274 кПа

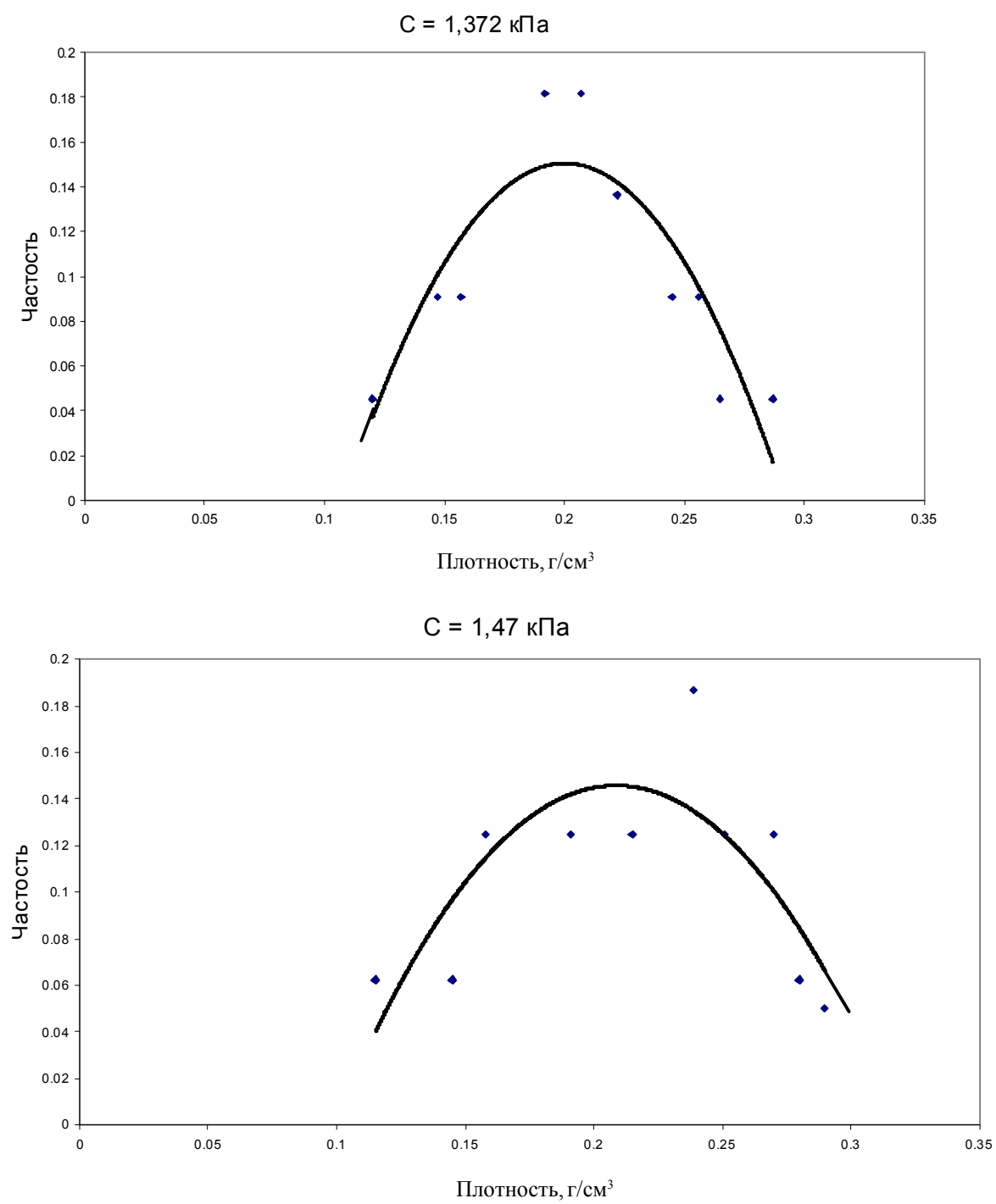


Рис. 5. Распределение величины плотности при сопротивлении сдвигу 1,372 и 1,47 кПа

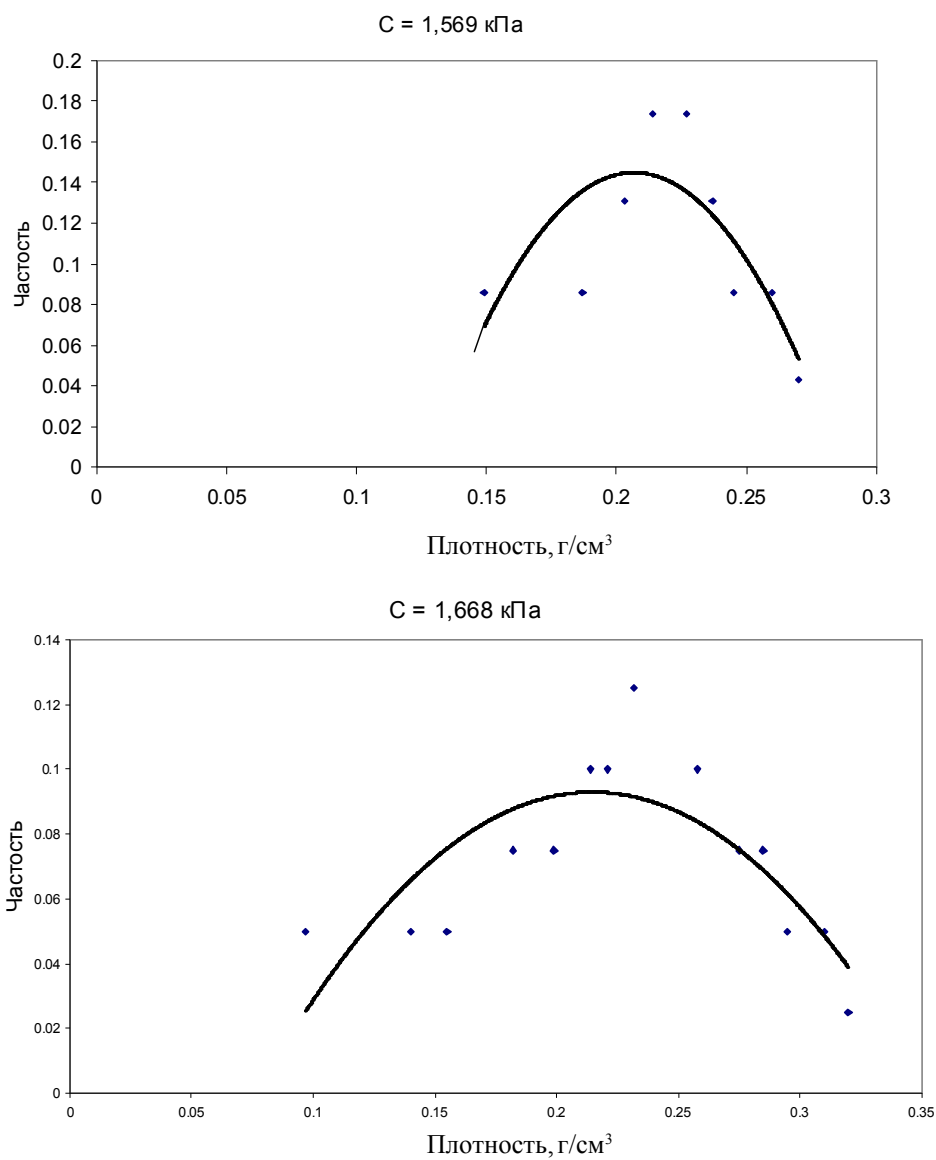


Рис. 6. Распределение величины плотности при сопротивлении сдвигу 1,569 и 1,668 кПа

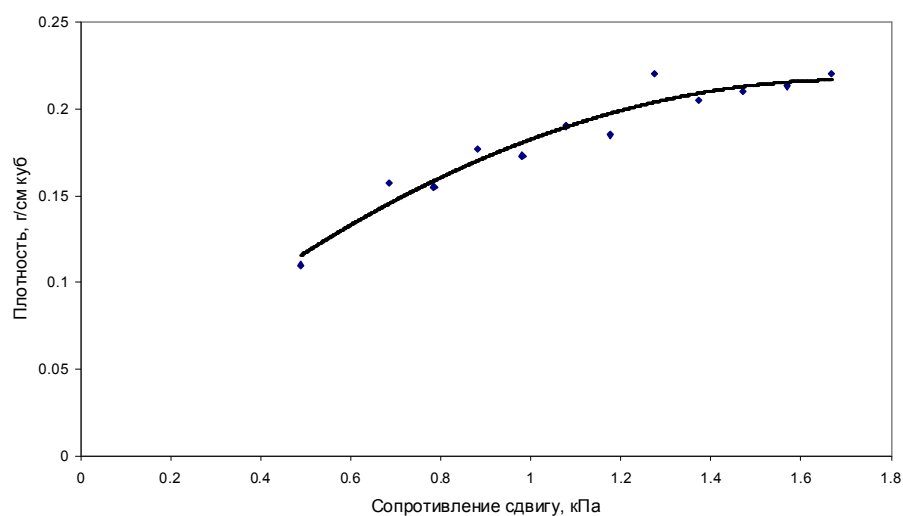


Рис. 7. Изменение значений сопротивления сдвигу и соответствующих им средних значений плотности

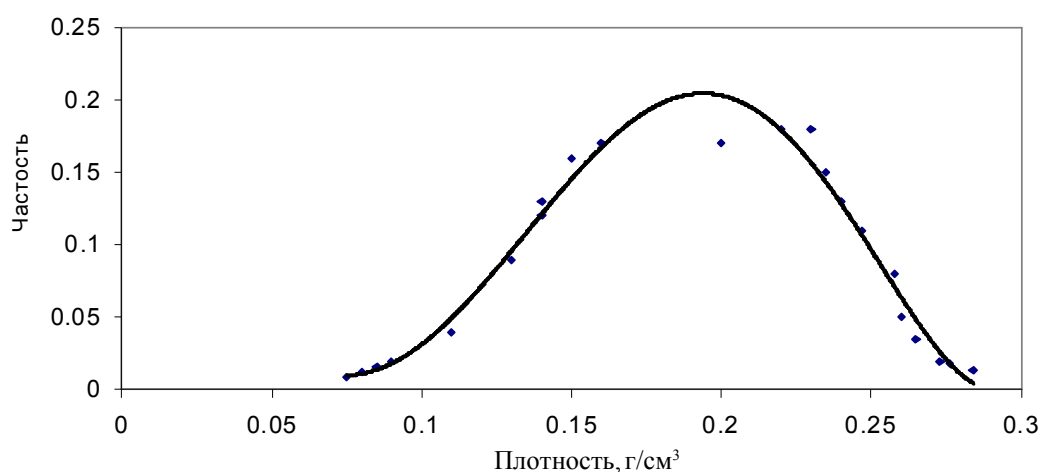


Рис. 8. Пример общего вида асимметричной одновершинной кривой Пирсона У1 типа

В табл. 1 приведены результаты расчетов коэффициента корреляции для приведенных выше диапазонов высоты снежного покрова. Из расчетов исключены только ряды переменных C и p , измеренные:

- 1) во влажном снеге по всей толще снежного покрова;
- 2) вблизи поверхности грунта, где плотность составляла до $0,57 \text{ г/см}^3$, а сопротивление сдвигу от $1,18$ до $1,96 \text{ кПа}$;
- 3) в приближенных к поверхности грунта слоях с сопротивлением сдвигу $4,41$ – $6,96 \text{ кПа}$, в то время когда в соседних верхних и нижних слоях величина C составляла от $27,0$ до $29,4 \text{ кПа}$ при одной и той же плотности, близкой к $0,4 \text{ г/см}^3$;
- 4) в слоях глубинной изморози.

Таблица 1

Коэффициенты корреляции по данным измерений в 63 профилях снежной толщи

Высота снежной толщи, см	Кол-во выборок	$R_{\text{эмп}}$			$R_{\text{кр}}$			$\frac{R_{\text{эмп}}}{R_{\text{кр}}}$		
		max	min	Сред.	max	min	Сред.	max	min	Сред.
50–100	24	0,956	0,577	0,827	0,9775	0,663	0,8694	0,999	0,800	0,951
110–140	17	0,918	0,452	0,744	0,9777	0,668	0,8256	0,996	0,639	0,895
150–350	12	0,982	0,564	0,799	0,9880	0,569	0,8192	0,999	0,909	0,975

Примечание. $R_{\text{эмп}}$ — коэффициент корреляции эмпирических рядов переменных; $R_{\text{кр}}$ — коэффициент корреляции аппроксимирующей кривой.

Корреляционное отношение $R_{\text{эмп}}/R_{\text{кр}}$ обычно применяют при использовании критерия статистической проверки гипотезы об общем виде функциональной зависимости, не требующего предварительного построения эмпирической линии регрессии, но позволяет проверять только линейный вид зависимости.

Кривые функций $C = f(p)$ и $p = f(C)$ приведены на рис. 9–11.

Затем для каждого диапазона снежной толщи по совокупности измерений и по общей выборке для каждого диапазона определены коэффициенты корреляции, приведенные в табл. 2 и 3.

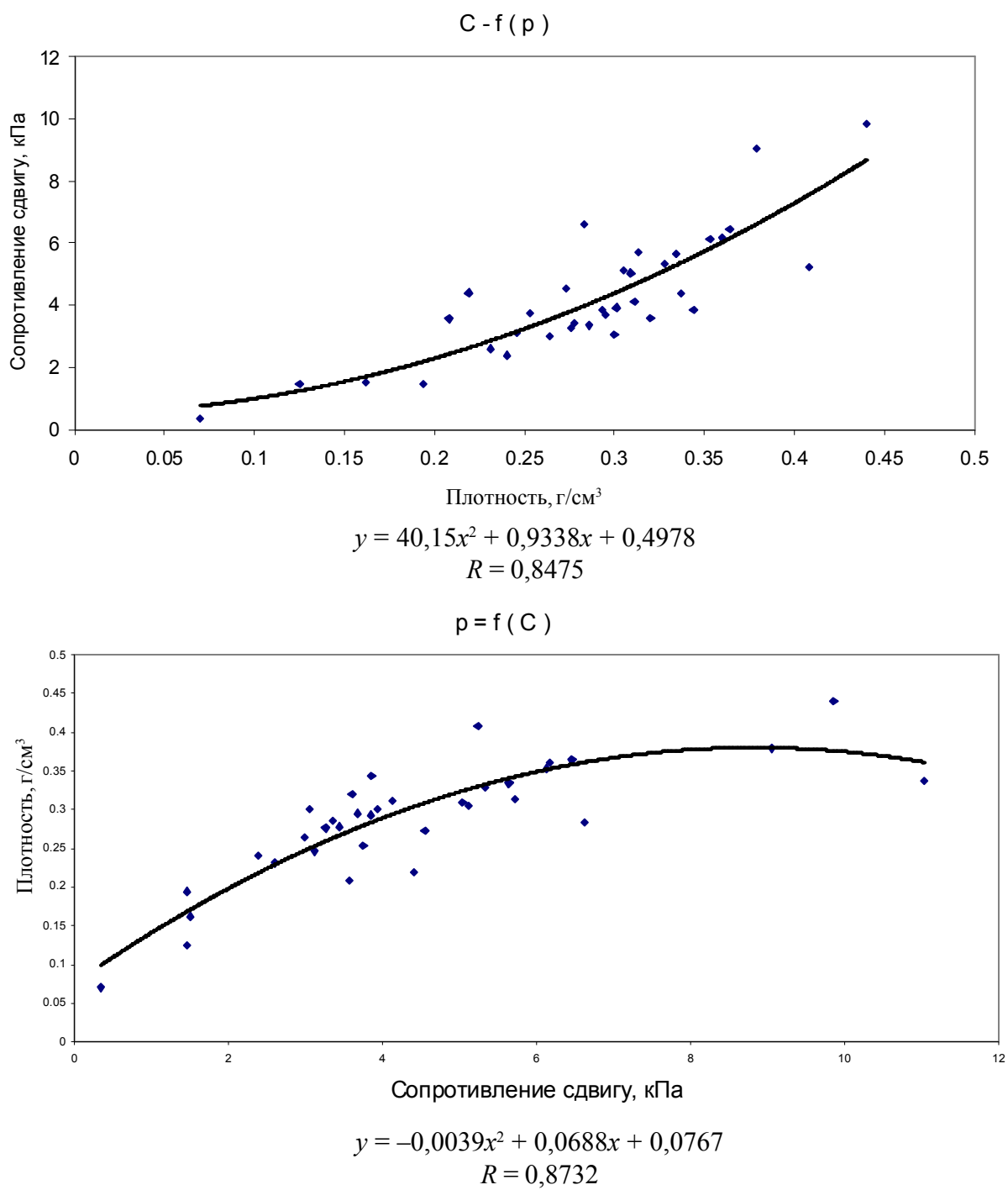


Рис. 9. Зависимость между плотностью и сопротивлением сдвигу по данным выборки из измерений в снежной толще высотой 50–100 см

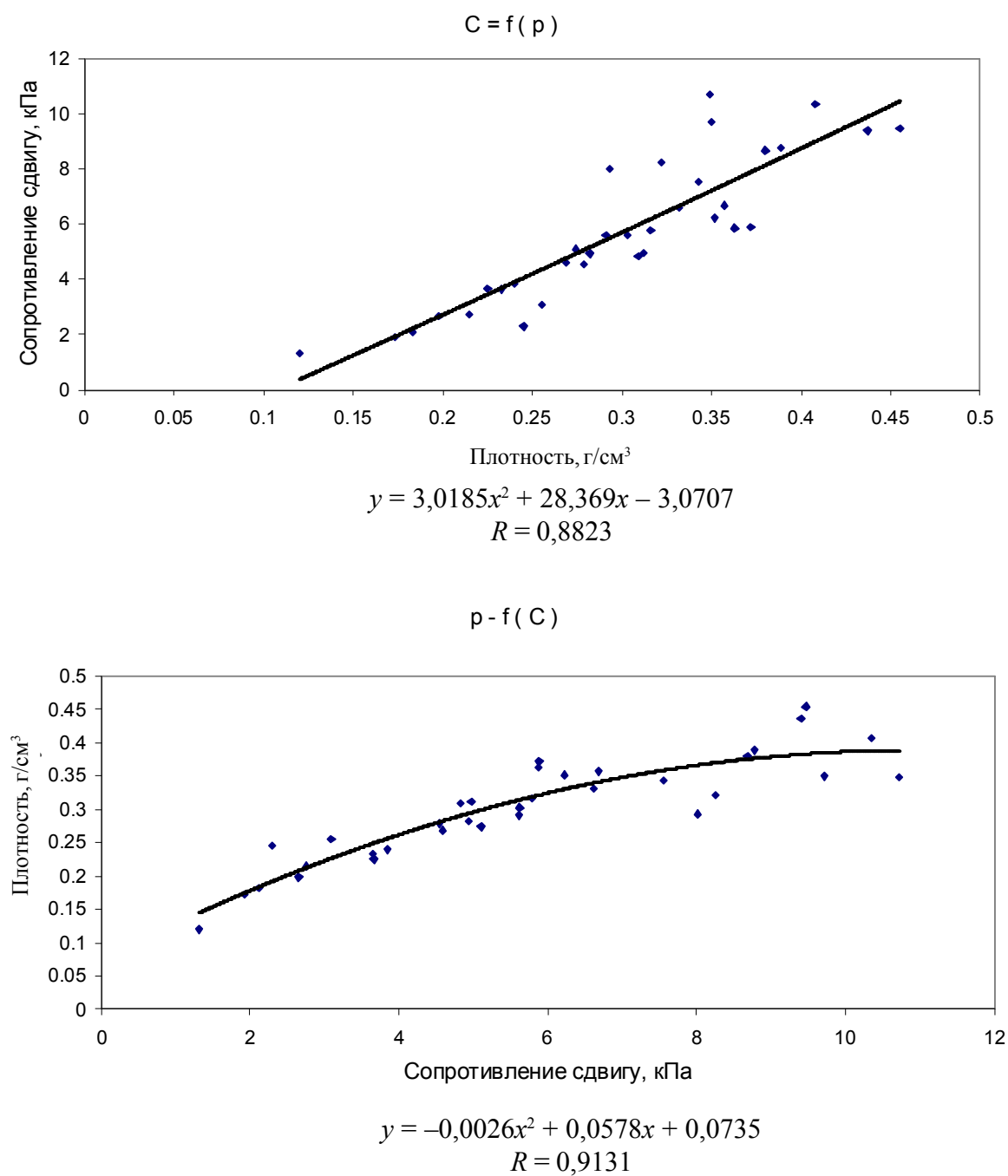


Рис. 10. Зависимость между плотностью и сопротивлением сдвигу по данным выборки из измерений в снежной толще высотой 110–140 см

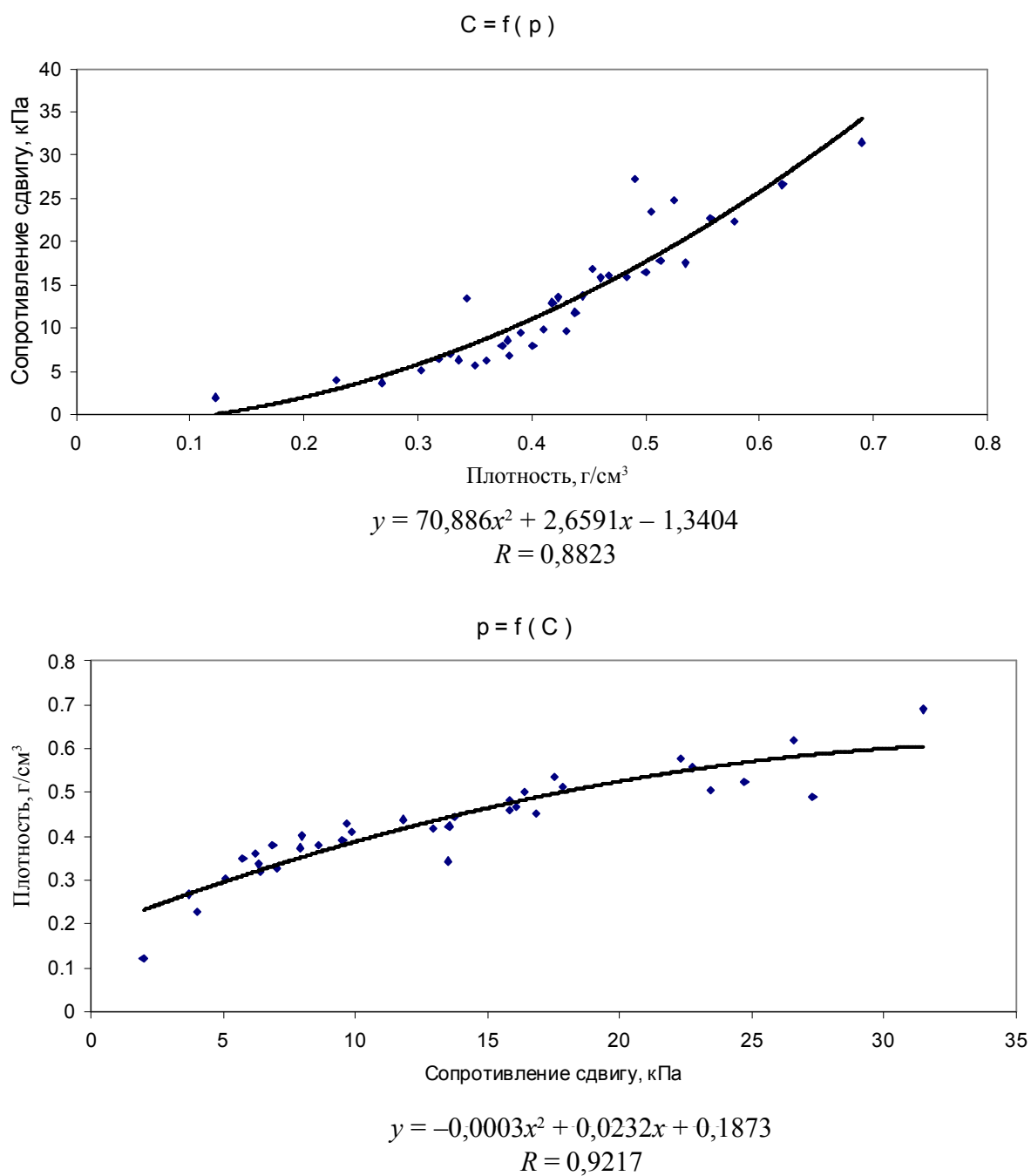


Рис. 11. Зависимость между плотностью и сопротивлением сдвигу по данным выборки из измерений в снежной толще высотой 150–350 см

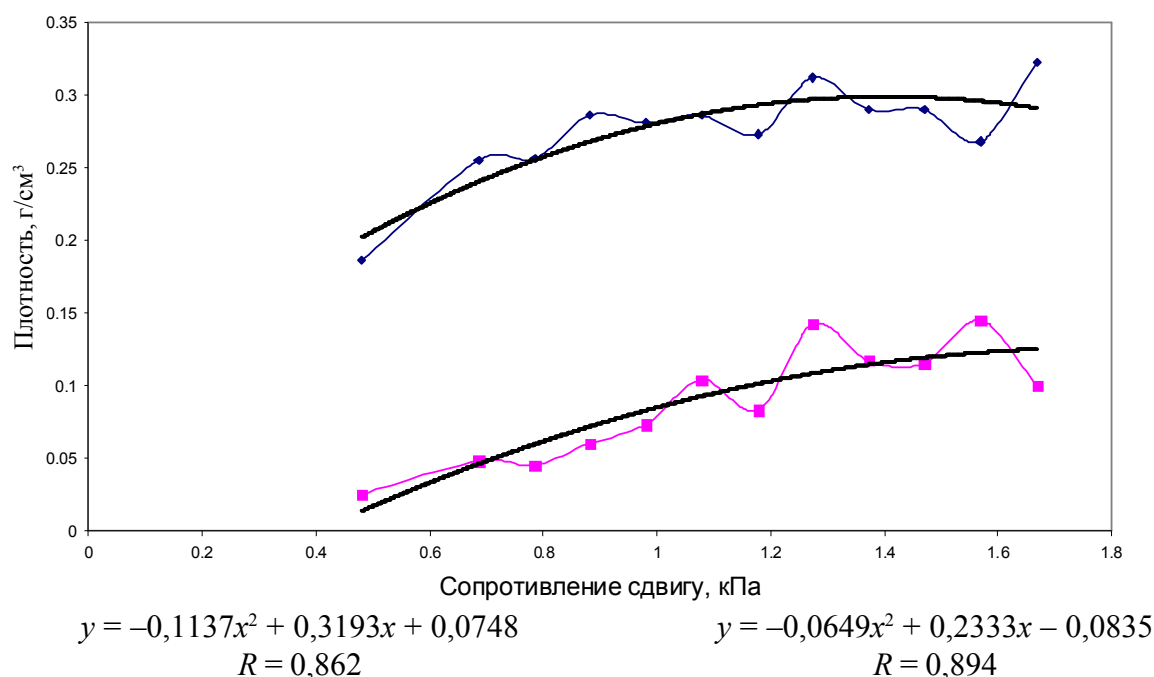


Рис. 12. Изменение величины сопротивления сдвигу, кПа, и соответствующих ему минимального и максимального значений плотности, г/см³

Таблица 2

Коэффициенты корреляции R для $C = f(\rho)$

Высота снежной толщи, см	По совокупности измерений R_1	По выборке R_2	По уравнению связи совокупности R_3	По уравнению связи выборки R_4	Достоверность аппроксимации по выборке R_4^2
50–100	0,6258	0,8288	0,6399	0,8475	0,7182
110–140	0,6046	0,8823	0,6542	0,8823	0,7785
150–350	0,7619	0,9062	0,7777	0,9249	0,8554

Таблица 3

Коэффициенты корреляции R для регрессионных кривых $\rho = f(C)$

Высота снежной толщи, см	По совокупности измерений R_1	По выборке R_2	По уравнению связи совокупности R_3	По уравнению связи выборки R_4	Достоверность аппроксимации по выборке R_4^2
50–100	0,5963	0,9255	0,6775	0,8732	0,7625
110–140	0,6042	0,8823	0,6563	0,9131	0,8338
150–350	0,7619	0,9062	0,8347	0,9217	0,8495

Из таблиц видно, что разброс величин коэффициента корреляции от 0,5963 (по совокупности измерений для снежной толщи $H_{\max} = 100$ см) до 0,9249 (по уравнению связи для выборки при $H_{\max} = 350$ см). Достоверность аппроксимации для криволинейной параболической зависимости второго порядка возрастает по мере роста высоты снежного покрова от 0,6341 до 0,8554 для всех регрессионных кривых.

Выполним статистическую оценку полученных коэффициентов корреляции на основе связи между плотностью и сопротивлением сдвигу, определенной криволинейной зависимостью в виде параболы второго порядка [2]. Оценку произведем для

уравнений совокупности и уравнений по выборке. Исходные данные для расчетов приведены в табл. 4.

Таблица 4

Данные для вычислений по статистической оценке

Диапазон высоты снежного покрова, см	Кол-во данных по уравнению совокупности n_1	Кол-во данных по выборке n_2	Коэф. корреляции R			
			по совокупности		по выборке	
			$C = f(p)$	$p = f(C)$	$C = f(p)$	$p = f(C)$
50–100	218	36	0,6062	0,6775	0,7903	0,8792
110–140	202	34	0,6542	0,6563	0,8823	0,9131
150–350	206	35	0,7777	0,8347	0,9249	0,9217

В расчетах принято число параметров в уравнении связи $m = 3$. Выражение $n - m$ определяет степени свободы вариации. Задача состоит в оценке достоверности полученного коэффициента корреляции и установлении пределов его случайных колебаний.

Находим стандартную ошибку коэффициента корреляции по формуле

$$\delta_i = \frac{1 - R^2}{\sqrt{n_i - m}}. \quad (3)$$

Затем вычислим фактическое нормированное отклонение $t_{\text{факт}}$:

$$t_{\text{факт}} = \frac{R}{\delta_i}. \quad (4)$$

По таблице t -распределения Стьюдента — Госсета устанавливаем величину $t_{\text{табл}}$ по числу степеней свободы $n_i - m$ и при 5 %-м доверительном уровне вероятности для одностороннего критерия $2Q$.

Расчет доверительных пределов случайных колебаний коэффициента корреляции выполняем по формуле

$$\varepsilon_i = t_{\text{табл}} \delta_i.$$

Результаты вычислений сведены в табл. 5.

Таблица 5

Вычисленные характеристики коэффициентов корреляции

Высота снежного покрова, см	Коэф. корреляции	Статистические характеристики				
		δ_i	$t_{\text{факт}}$	$t_{\text{табл}}$	ε	$\varepsilon_{\text{сред}}$
50–100	0,6062	0,0431	14,065	1,98	0,0853	0,094
	0,6775	0,0369	18,36	1,98	0,0730	
	0,7903	0,0653	12,1026	2,034	0,1328	
	0,8792	0,0413	21,1428	2,034	0,0840	
110–140	0,6542	0,0405	16,1531	1,98	0,0802	0,076
	0,6563	0,0404	16,2450	1,98	0,080	
	0,8823	0,0398	22,1683	2,048	0,0815	
	0,9131	0,0299	33,1542	2,048	0,0612	
150–350	0,7777	0,0277	28,076	1,98	0,0548	0,051
	0,8347	0,0213	39,188	1,98	0,0422	
	0,9249	0,0255	36,27	2,036	0,0519	
	0,9217	0,0266	34,651	2,036	0,0542	

Из таблицы видно, что величина стандартной ошибки δ_i уменьшается с ростом высоты снежного покрова и модуля коэффициента корреляции почти в 2–2,5 раза. Фактические нормированные отклонения $t_{\text{факт}}$ почти для всех коэффициентов корреля-

ции на порядок больше $t_{\text{табл}}$ и можно говорить об удовлетворительной связи между измеренными плотностью и сопротивлением сдвигу, представленной параболическим уравнением. Случайные колебания величины коэффициента корреляции ε , определяющие доверительный интервал колебаний R , только в единственном случае могут повлиять на первую значащую цифру модуля R . В целом прослеживается снижение ε почти в 2 раза с увеличением высоты снежного покрова.

К таким вопросам регрессионного анализа, как построение конкретного вида зависимости между переменными, различные оценки точности этой зависимости относятся также вопросы, связанные с исследованием степени тесноты связи между переменными, проведение дисперсионного анализа.

Из курса теории вероятностей известно, что дисперсия D дискретной случайной величины — это математическое ожидание квадратов отклонения ее значения от ее математического ожидания. Дисперсия зависит только от длины интервала между величинами и является возрастающей функцией длины. Чем больше интервал значений, т.е. чем больше рассеяны значения величины, тем больше дисперсия. Дисперсия играет роль меры разбросанности значений случайной величины около математического ожидания.

Применим метод дисперсионного анализа для оценки существенности отклонения от линейной регрессии. Он основан на сопоставлении зависимостей линейного и криволинейного характера [2].

Выдвинем гипотезу о том, что криволинейное уравнение не дает лучшей меры зависимости по сравнению с линейной. Задача решается с помощью дисперсионного анализа, для выполнения которого необходимы линейное и криволинейное (параболическое) уравнения:

$$Y = Ax + B, \quad (5)$$

$$Y = Cx^2 + Dx + E, \quad (6)$$

а также вычисленные статистические параметры $\Sigma y =$; $\Sigma x =$; $\Sigma y^2 =$; $\Sigma xy =$; $\Sigma x^2 y =$; $\Sigma y^2 x =$. Примем уровень вероятности $P = 0,05$.

По известным формулам рассчитываем остаточные суммы квадратов отклонений по обоим уравнениям.

Для линейного уравнения

$$Q'_{\text{ост}} = \Sigma y^2 - B \Sigma y - A \Sigma xy. \quad (7)$$

Для параболического уравнения

$$Q''_{\text{ост}} = \Sigma y^2 - E \Sigma y - D \Sigma xy - C \Sigma x^2 y. \quad (8)$$

Определяем разность между двумя суммами квадратов отклонений

$$Q_{\text{разн}} = Q'_{\text{ост}} - Q''_{\text{ост}}. \quad (9)$$

Для определения остаточных дисперсий определяем степени свободы вариации.

Число степеней свободы в каждом уравнении связи $k = n - m$, где n — количество наблюдений в выборке; m — количество параметров в уравнении связи.

Для получения уравнений связи функций $C = f(p)$ и $p = f(C)$ трех диапазонов высоты снежного покрова 50–100, 110–140 и 150–350 см значения n приведены в табл. 4. Для линейных уравнений $m_1 = 2$, параболических $m_2 = 3$. Для определения дисперсии разности сумм квадратов отклонений, рассчитанных по уравнениям (6)–(9), степень свободы получаем как разность между степенями свободы линейного и параболического уравнений

$$k_{\text{разн}} = (n - m_1) - (n - m_2) = 1. \quad (10)$$

По известным $Q'_{\text{ост}}$ и $Q''_{\text{ост}}$ вычисляем остаточную дисперсию по параболическому уравнению $D'_{\text{ост}}$ и дисперсию разности $D_{\text{разн}}$.

$$D''_{\text{ост}} = \frac{Q''_{\text{ост}}}{n - m_2} \text{ и } D_{\text{разн}} = \frac{Q'_{\text{ост}} - Q''_{\text{ост}}}{(n - m_1) - (n - m_2)}. \quad (11)$$

Отношение дисперсий $\frac{D_{\text{разн}}}{D''_{\text{ост}}}$ дает величину фактического распределения эмпирической зависимости $F_{\text{факт}}$. По таблице распределения $F(v_1, v_2)$ для 5 %-го уровня находим $F_{\text{табл}}$. Для дисперсии разности число степеней свободы $v_1 = 1$, для v_2 определяем по расчету. Результаты расчетов приведены в табл. 6.

Таблица 6

Результаты расчетов к дисперсионному анализу

Высота снежного покрова, см	Число измерений по выборке	$C = f(p)$				$p = f(C)$			
		$D_{\text{разн}}$	$D''_{\text{ост}}$	$F_{\text{факт}}$	$F_{\text{табл}}$	$D_{\text{разн}}$	$D''_{\text{ост}}$	$F_{\text{факт}}$	$F_{\text{табл}}$
1	2	3	4	5	6	7	8	9	10
50–100	36	4,115	2,039	2,018	4,14	0,0304	0,0014	21,7	4,14
110–140	34	0,309	1,616	0,019	4,16	0,0349	0,00031	112,6	4,16
150–350	35	67,619	8,895	7,60	4,12	0,0625	0,00007	833	4,12

Из таблицы видно, что:

1) в тех случаях, когда $F_{\text{факт}} > F_{\text{табл}}$ (или $F_{\text{факт}} \gg F_{\text{табл}}$) гипотеза о криволинейной (параболической) зависимости между переменными подтверждена (все регрессионные кривые функции $p = f(C)$, столбцы 5 и 9);

2) переменные C и p не всегда связаны точной линейной зависимостью и часто вторая формула имеет самостоятельное значение (столбцы 4 и 8, 3 и 7);

3) различия в оценке F -распределения переменных, возможно, зависит от высоты снежного покрова, но не от количества измерений. При общем количестве переменных в каждом диапазоне чуть более 200 их изменение составило $208 \pm 4 \%$;

4) в первых двух строках столбца 5 таблицы значения $F_{\text{факт}} < F_{\text{табл}}$, что свидетельствует о небольшом разбросе переменных и возможном описании зависимости линейным уравнением. Для $F_{\text{факт}} = 0,019$ это хорошо видно на рис. 9 — $C = f(p)$.

Статистическая проверка гипотезы о принятом виде функциональной зависимости возможна и по одной из ранее упомянутых в статье схем. А именно по схеме 3, учитывающей зависимость переменной от совокупности неконтролируемых случайных факторов. Находясь в связи друг с другом, значения C и p заметно варьируют при послойном измерении и в различных профилях снежной толщи. Проверка позволяет либо подтвердить приемлемость исследуемого вида зависимости, либо отвергнуть. Разработаны статистические критерии проверки, рекомендуемые также для широкого спектра криволинейных зависимостей [1].

Выполним проверку для диапазона высоты снежного покрова 50–100 см, в котором $F_{\text{факт}} < F_{\text{табл}}$ (табл. 6, столбцы 5 и 6). Проверку сделаем по критерию, который предполагает использование конкретного вида эмпирической регрессионной зависимости $\hat{Y} = f(x)$ и основан на вычислении параметра v^2 по уравнению

$$v^2 = \frac{(n - k) \sum_{i=1}^k m_i [\bar{y}_i - \overline{Y(x_i^0)}]^2}{(k - g) \sum_{i=1}^k \sum_{j=1}^{m_i} [y_{ij} - \bar{y}_i]^2}. \quad (12)$$

Если числитель уравнения (12) ненамного больше знаменателя или примерно равен ему, принятая зависимость считается приемлемой.

В случае правильного выбора регрессионной зависимости величина u^2 должна быть либо равна, либо чуть меньше $F(v_1, v_2)$ — распределения с числом степеней свободы $v_1 = k - g$ и знаменателя $v_2 = n - k$. В формуле n — общее количество измерений; k — количество интервалов, на которые разбит диапазон независимой переменной x ; g — количество параметров уравнения кривой; m_i — количество точек в i -м интервале. Имеем для рассматриваемых переменных C и p уравнение зависимости в интервале высот снега 50–100 см.

$$y = 40,151x^2 + 0,9338x + 0,4978. \quad (13)$$

В уравнении (12) числитель определяет меру разброса осредненных опытных измерений около эмпирической регрессионной кривой $\hat{Y}(x)$. Знаменатель характеризует меру разброса опытных измерений около своих частных средних $\bar{y}_i$ т.е. меру, не зависящую от выбранного вида регрессионной кривой. Численные значения параметров $n = 216$, $k = 36$, $g = 3$, $m_i = 6$ и степени свободы $v_1 = 33$ и $v_2 = 1,6$. Числитель и знаменатель в уравнении считаются независимыми статистическими оценками одной и той же теоретической дисперсии σ^2 . Предварительно вычислены статистические параметры, входящие в уравнение (12), и значения полученной регрессионной функции по всему ряду точек (что создавало неудобство вычислительного характера).

Для числителя получено $\sum_{i=1}^k m_i [\bar{y}_i - \overline{Y(x_i)}]^2 = 213,5327$, для знаменателя $\sum_{i=1}^k \sum_{j=1}^{m_i} [y_{ij} - \bar{y}_i]^2 = 706,8886$, в результате $u^2 = 0,302$. Числитель оказался меньше знаменателя, поэтому выбранный вид регрессионной зависимости можно считать приемлемым.

Более точно приемлемость выбранной зависимости можно определить по таблице $F(v_1, v_2)$ — распределения при достаточно большом уровне значимости $P = 0,01$. Полученные v_1 и v_2 дают $F_{\text{табл}} = 1,79$. Поскольку $u^2 < F_{\text{табл}}$, то выбранный вид регрессионной зависимости не противоречит измеренным данным. С вероятностью ошибиться в одном случае из ста можно утверждать, что связь между плотностью и сопротивлением сдвигу в слежавшемся снежном покрове достоверна.

Предпринятые ранее немногие попытки анализа связи между плотностью и сопротивлением сдвигу в снежной массе свидетельствуют о том, что физико-механические свойства снега носят не функциональный, а вероятностный характер. Представленный в статье профессионально-технический анализ приводит к возможности реализации следующей схемы исследования:

1) сбор данных измерений с применением различных приемов первичной статистической обработки (исключение резко выделяющихся результатов, проверка их вероятностной зависимости, проверка однородности выборки);

2) рассмотрение природы исходных данных, их зависимости от условий формирования и залегания снежного покрова для установления типа статистической зависимости;

3) проверка степени отклонения распределения зависимой переменной от нормального;

4) предварительный анализ полученного корреляционного поля. Возможность в случае большого количества данных разбиения диапазона изменения переменных на интервалы;

5) оценка степени тесноты связи (определение коэффициентов корреляции, корреляционных отношений);

6) установление общего вида зависимости (линейная, параболическая, степенная и т.д.).

Корреляционные связи между рассмотренными характеристиками определяются их физической взаимообусловленностью. Специалисты рекомендуют при переходе от коррелированности между величинами к установлению объективной связи между ними в первую очередь исследовать степень отклонения распределения рассматриваемых переменных от нормального, но только в случае одновершинности распределения и приближенной его симметричности. Для снежного покрова дополнительно необходим по возможности максимально корректный анализ общих факторов (физических величин), изменение которых вызывает определенную тенденцию в одновременном изменении исследуемых величин.

Выполненный анализ позволяет сделать следующие выводы:

1) между такими физико-механическими свойствами снежного покрова, как плотность и сопротивление сдвигу существуют связи, выявить которые можно с помощью аксиом и постулатов теории вероятностей;

2) рассмотренные зависимости между C и p позволяют относить их к величинам случайным дискретным. Зависимости могут быть установлены и проверены методами оценки с помощью статистических критериев и дисперсионного анализа;

3) анализ обработки регрессионных зависимостей показывает, что в некоторых случаях характер обратных зависимостей может иметь самостоятельное значение;

4) в целом гипотеза о параболическом виде регрессионной зависимости между C и p не противоречит данным опытных измерений;

5) максимальная вероятность значений сопротивления сдвигу растет с ростом плотности слоя снега.

Библиографический список

1. Айвазян С.А. Статистические методы исследований. М.: Металлургия, 1968. 227 с.
2. Брукс К., Карузерс Н. Применение статистических методов в метеорологии. Л.: Гидрометеиздат, 1963. 416 с.
3. Вериго М.Ф. Методическое пособие по применению математической статистики в обработке опытных данных. Новосибирск: НИИЖТ, 1964. 114 с.
4. Марин Ю.А. О распределении плотности и сопротивления сдвигу по профилю снежной толщи // Железные и автомобильные дороги в условиях Сибири. Новосибирск, 2009. С. 91–124.
5. Зельдович Я.В., Мышкис А.Д. Элементы прикладной математики. М.: Наука, 1967. 646 с.
6. Гнеденко Б.В. Курс теории вероятностей. М.: Наука, 1969. 400 с.

Y.A. Marin. Correlation-Regressive Analysis of Density and Shift Resistance Interrelation in a Formed Depth of Snow Cover.

Results of statistic analysis of physical and mechanical snow properties interrelation are given, that is of shift and density resistance. These characteristics were measured in pairs in 10-cm layers of snow cover during 13 winters from December to April in a mountain avalanche-prone area of South Sakhalin. Density values distribution curves, measured for each of 12 values of shift resistance in a range of 0,49 kPa to 1,49 kPa are obtained. Statistic assessment of correlation coefficients is done. Dependence of line and parabolic character are compared by the dispersion analysis. Hypothesis of obtained functional dependence is tested. Research plan of possible realization and conclusions on the analysis results are proposed.

Key words: *physical and mechanical snow properties, snow avalanches, shift resistance, density, statistic analysis, correlation coefficient.*